B.T.

Industria y Comercio

SUPERINTENDENCIA

PCT

SOLICITUD

PATENTE DE INVENCIÓN

- PCT ENTRADA EN FASE NACIONAL -

21 EXPEDIENTE Nº 04-178803

54 TÍTULO "TOXINAS DE CRY3A MODIFICADAS Y SECUENCIAS DE ACIDO NUCLEIDO QUE LAS MODIFICAN"

51 CLASIFICACIÓN INTERNACIONAL C12N15/31, 15/82, C07K 14/325 A61K31/713

71 SOLICITANTE SYNGENTA PARTICIPATIONS AG

DOMICILIO Schwarzwaldallee 215 CH-4058 Basilea, SUIZA

74 APODERADO DRA MARGARITA CASTELLANOS ABONDANO

22 BOGOTÁ, D C 27 FEB 2004

(FORMA P 10)

PETITORIO

Industria y Com
SUPERINTENDEN

04 017803 00

SUPERINDUSTRIA Y COMERCIO
Radicacion : 04017803 00000000 Folios: 405
Fecha (AMD): 2004-02-27 16:26:15
Tramite : 011 PCT-PATENT 379 FASE NACI 411 PRESENTAC
Dependencia: 2020 DIVISION DE NUEVAS CREACIONES

FORMULARIO UNICO DE SOLICITUD DE PATENTE
PCT
ENTRADA EN FASE NACIONAL 31-Marzo-2004

SOLICITUD DE:

(1)

[X] Patente de Invención. [ ] Patente de Modelo de Utilidad.

[ ] Capítulo I [X] Capítulo II

(2) **SOLICITANTE (71)**	Nombre: **SYNGENTA PARTICIPATIONS AG** Dirección: **Schwarzwaldallee 215, CH-4058 Basilea - SUIZA** Nacionalidad o Domicilio: **Basilea - SUIZA** Lugar de Constitución: **SUIZA** Teléfono: Fax: E-mail:	**IDENTIFICACIÓN** C.C. [ ] NIT [ ] C.E. [ ] Otro [ ] Cual ____ Número ____
(3) **REPRESENTANTE O APODERADO (74)**	Nombre: **MARGARITA CASTELLANOS A.** Dirección: **CARRERA 9 No. 93-09** Teléfono: **610 74 80** Fax: **610 07 06** E-mail: **acastell@impsat.net.co**	**IDENTIFICACIÓN** C.C. [X] NIT [ ] C.E. [ ] Otro [ ] Número **38'779.654** TPA **70.819**
(4) **INVENTOR (ES) (72)**	Nombre: **ERIC CHEN y CHERYL STACY** Dirección: **para los dos inventores es: 3054 Cornwallis Road, Research Triangle Park, NC 27709, ESTADOS UNIDOS DE AMERICA** Nacionalidad o Domicilio **Estadounidense**	

(5) **PCT (86)**

Solicitud Internacional No. **PCT/US01/29613** Fecha **21/09/2001**

Publicación Internacional No. **WO 02/037632 A3** Fecha **10/5/02**

(6) **Título (54)** **"TOXINAS DE CRY3A MODIFICADAS Y SECUENCIAS DE ACIDO NUCLEICO QUE LAS MODIFICAN"**

(7) **Clasificación Internacional (51)** C12N15/31; 15/82; C07K14/325; A61K31/713

(8) **Prioridad** SI [X] NO [ ]

(33) País de Origen	(32) Fecha	(31) Número de Solicitud
US	31/08/2001	60/316421

(9) **Comprobante de pago No.** **04-14.976** **Fecha** **Febrero 25, 2004**

Instrucciones para el diligenciamiento del presente formulario, ver reverso de esta página

(10) Anexos

- [X] Comprobante de pago de la tasa de presentación de la solicitud.
- [ ] Documento que acredita la existencia y representación legal cuando el solicitante sea persona jurídica.
- [ ] Poderes, si fuere el caso.
- [ ] Documento de cesión del inventor al solicitante o a su causante.
- [ ] Declaraciones.
- [X] Copia de la solicitud internacional (Resumen, descripción, reivindicaciones, dibujos).
- [X] Traducción de la solicitud internacional al castellano (Resumen, descripción, reivindicaciones, dibujos).
- [X] Copia del examen preliminar internacional efectuado por la Autoridad Internacional de Examen Preliminar (IPEA).
- [ ] Traducción del examen preliminar internacional efectuado por la IPEA.
- [ ] De ser el caso, copia del contrato de acceso.
- [ ] De ser el caso, documento que acredite la licencia o autorización de uso de conocimientos tradicionales de las comunidades indígenas.
- [ ] De ser el caso, certificado de depósito del material biológico.
- [ ] Arte final 12x12
- [ ] De ser el caso, información sobre otras solicitudes de patente o títulos obtenidos en el extranjero por el mismo titular o su causante, relacionadas parcial o totalmente con la invención de esta solicitud.
- [ ] De ser el caso, modificaciones efectuadas a la solicitud internacional.
- [X] Otros Informe de busqueda internacional, copia certificada del documento de prioridad

(11) FIGURA CARACTERISTICA:

(12) Solicito la concesión de la patente,

27 FEB. 2004

NOMBRE: MARGARITA CASTELLANOS ABONDANO

FIRMA:

C.C. No: 39.779.654 Usaquén - TPA 70.819 C.S.J.

Instrucciones para el diligenciamiento del presente formulario, ver reverso de esta página. - MCA/gcb. -

3

SUPERINDUSTRIA Y COMERCIO
Radicacion : 04017603 00000000 Folios: 405
Fecha (AMD): 2004-02-27 16:28:15
Tramite : 011 PCT-PATENT 379/FASE NACI 411 PRESENTAC
Dependencia: 2020 DIVISION DE NUEVAS CREACIONES

NIT : 800176089-2

RECIBO OFICIAL DE CAJA : 04 - 14,976
FECHA : FEBRERO 25 DE 2004

***** C O N S I G N A C I O N *****

DEPOSITANTE	TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PGO	Vr. PAGO
HILDA MARIN	CONSIGNACION	BANCO POPULAR	050-00140-6	8780760	25/02/2004	8,044,000.00

***** C O N C E P T O *****

CANT.	RENTISTICO	CONCEPTO	TOTAL CONCEPTO
1	50005-01-01 SOLICITUDES	1 TRAMITES DE SOL. DE PATENTE DE IN	422,000.00
		TOTAL : $	422,000.00

SON: CUATROCIENTOS VEINTIDOS MIL PESOS

RESPONSABLE : _____

RECIBO DE CAJA APLICADO AL EXPEDIENTE No._____

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
6 March 2003 (06.03.2003)

PCT

(10) International Publication Number
WO 03/018810 A2

(51) International Patent Classification[7]: C12N 15/31, 15/82, C07K 14/325, A01H 5/00

(21) International Application Number: PCT/EP02/09789

(22) International Filing Date: 2 September 2002 (02.09.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/316,421 31 August 2001 (31.08.2001) US

(71) Applicant *(for all designated States except US)*: SYNGENTA PARTICIPATIONS AG [CH/CH]; Schwarzwaldallee 215, CH-4058 Basel (CH).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: CHEN, Eric [US/US]; Syngenta Biotechnology, Inc., 3054 Cornwallis Road, Research Triangle Park, NC 27709 (US). STACY, Cheryl [US/US]; Syngenta Biotechnology, Inc., 3054 Cornwallis Road, Research Triangle Park, NC 27709 (US).

(74) Agent: BASTIAN, Werner; 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, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, 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.

WO 03/018810 A2

(54) Title: MODIFIED CRY3A TOXINS AND NUCLEIC ACID SEQUENCES CODING THEREFOR

(57) Abstract: Compositions and methods for controlling plant pests are disclosed. In particular, novel nucleic acid sequences encoding modified Cry3A toxins having increased toxicity to corn rootworm are provided. By inserting a protease recognition site, such as cathepsin G, that is recognized by a gut protease of a target insect in at least one position of a Cry3A toxin a modified Cry3A toxin having significantly greater toxicity, particularly to western and northern corn rootworm is designed. Further, a method of making the modified Cry3A toxins and methods of using the modified Cry3Anucleic acid sequences, for example in microorganisms to control insects or in transgenic plants to confer protection from insect damage, and a method of using the modified Cry3A toxins, and compositions andformulations comprising the modified Cry3A toxins, for example applying the modified Cry3A toxins or compositions or formulations to insect-infested areas, or to prophylactically treat insect-susceptible areas or plants to confer protection against the insect pests are disclosed.

PATENT COOPERATION TREATY

MAY 1 6 2003

From the INTERNATIONAL BUREAU

PCT

NOTICE INFORMING THE APPLICANT OF THE COMMUNICATION OF THE INTERNATIONAL APPLICATION TO THE DESIGNATED OFFICES

(PCT Rule 47.1(c), first sentence)

To:

BASTIAN, Werner
Syngenta Participations AG
Intellectual Property
P.O. Box
CH-4002 Basel
SUISSE

Date of mailing *(day/month/year)*	
06 March 2003 (06.03.03)	
Applicant's or agent's file reference	IMPORTANT NOTICE
S-60065A	

International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/EP02/009789	02 September 2002 (02.09.02)	31 August 2001 (31.08.01)

Applicant

SYNGENTA PARTICIPATIONS AG

1. Notice is hereby given that the International Bureau has communicated, as provided in Article 20, the international application to the following designated Offices on the date indicated above as the date of mailing of this notice:

   KP, KR, US

   In accordance with Rule 47.1(c), third sentence, those Offices will accept the present notice as conclusive evidence that the communication of the international application has duly taken place on the date of mailing indicated above and no copy of the international application is required to be furnished by the applicant to the designated Office(s).

2. The following designated Offices have waived the requirement for such a communication at this time:

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

   The communication will be made to those Offices only upon their request. Furthermore, those Offices do not require the applicant to furnish a copy of the international application (Rule 49.1(a-*bis*)).

3. Enclosed with this notice is a copy of the international application as published by the International Bureau on 06 March 2003 (06.03.03) under No. WO 03/018810.

4. **TIME LIMITS for filing a demand for international preliminary examination and for entry into the national phase**

   The applicable time limit for entering the national phase will, subject to what is said in the following paragraph, be **30 MONTHS** from the priority date, not only in respect of any elected Office if a demand for international preliminary examination is filed before the expiration of 19 months from the priority date, but also in respect of any designated Office, in the absence of filing of such demand, where Article 22(1) as modified with effect from 1 April 2002 applies in respect of that designated Office. For further details, see *PCT Gazette* No. 44/2001 of 1 November 2001, pages 19926, 19932 and 19934, as well as the *PCT Newsletter*, October and November 2001 and February 2002 issues.

   In practice, time limits other than the 30-month time limit will continue to apply, for various periods of time, in respect of certain designated or elected Offices. For regular updates on the applicable time limits (20, 21, 30 or 31 months, or other time limit), Office by Office, refer to the *PCT Gazette*, the *PCT Newsletter* and the *PCT Applicant's Guide*, Volume II, National Chapters, all available from WIPO's Internet site, at http://www.wipo.int/pct/en/index.html.

   For filing a demand for international preliminary examination, see the *PCT Applicant's Guide*, Volume I/A, Chapter IX. Only an applicant who is a national or resident of a PCT Contracting State which is bound by Chapter II has the right to file a demand for international preliminary examination (at present, all PCT Contracting States are bound by Chapter II).

   It is the applicant's sole responsibility to monitor all these time limits.

The International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland

Facsimile No.(41-22) 740.14.35

Authorized officer

Judith Zahra

Telephone No.(41-22) 338.91.11

Modified Cry3A Toxins and Nucleic Acid Sequences Coding Therefor

The present invention relates to the fields of protein engineering, plant molecular biology and pest control. More particularly, the present invention relates to novel modified Cry3A toxins and nucleic acid sequences whose expression results in the modified Cry3A toxins, and methods of making and methods of using the modified Cry3A toxins and corresponding nucleic acid sequences to control insects.

Species of corn rootworm are considered to be the most destructive corn pests. In the United States the three important species are *Diabrotica virgifera virgifera*, the western corn rootworm; *D. longicornis barberi*, the northern corn rootworm and *D. undecimpunctata howardi*, the southern corn rootworm. Only western and northern corn rootworms are considered primary pests of corn in the US Corn Belt. Corn rootworm larvae cause the most substantial plant damage by feeding almost exclusively on corn roots. This injury has been shown to increase plant lodging, to reduce grain yield and vegetative yield as well as alter the nutrient content of the grain. Larval feeding also causes indirect effects on maize by opening avenues through the roots for bacterial and fungal infections which lead to root and stalk rot diseases. Adult corn rootworms are active in cornfields in late summer where they feed on ears, silks and pollen, interfering with normal pollination.

Corn rootworms are mainly controlled by intensive applications of chemical pesticides, which are active through inhibition of insect growth, prevention of insect feeding or reproduction, or cause death. Good corn rootworm control can thus be reached, but these chemicals can sometimes also affect other, beneficial organisms. Another problem resulting from the wide use of chemical pesticides is the appearance of resistant insect varieties. Yet another problem is due to the fact that corn rootworm larvae feed underground thus making it difficult to apply rescue treatments of insecticides. Therefore, most insecticide applications are made prophylactically at the time of planting. This practice results in a large environmental burden. This has been partially alleviated by various farm management practices, but there is an increasing need for alternative pest control mechanisms.

Biological pest control agents, such as *Bacillus thuringiensis* (Bt) strains expressing pesticidal toxins like δ-endotoxins, have also been applied to crop plants with satisfactory results against

primarily lepidopteran insect pests. The δ-endotoxins are proteins held within a crystalline matrix that are known to possess insecticidal activity when ingested by certain insects. The various δ-endotoxins have been classified based upon their spectrum of activity and sequence homology. Prior to 1990, the major classes were defined by their spectrum of activity with the Cry1 proteins active against Lepidoptera (moths and butterflies), Cry2 proteins active against both Lepidoptera and Diptera (flies and mosquitoes), Cry3 proteins active against Coleoptera (beetles) and Cry4 proteins active against Diptera (Hofte and Whitely, 1989, Microbiol. Rev. 53:242-255). Recently a new nomenclature was developed which systematically classifies the Cry proteins based on amino acid sequence homology rather than insect target specificities (Crickmore *et al.* 1998, Microbiol. Molec. Biol. Rev. 62:807-813).

The spectrum of insecticidal activity of an individual δ-endotoxin from Bt is quite narrow, with a given δ-endotoxin being active against only a few species within an Order. For instance, the Cry3A protein is known to be very toxic to the Colorado potato beetle, *Leptinotarsa decemlineata*, but has very little or no toxicity to related beetles in the genus *Diabrotica* (Johnson *et al.*, 1993, J. Econ. Entomol. 86:330-333). According to Slaney *et al.* (1992, Insect Biochem. Molec. Biol. 22:9-18) the Cry3A protein is at least 2000 times less toxic to southern corn rootworm larvae than to the Colorado potato beetle. It is also known that Cry3A has little or no toxicity to the western corn rootworm.

Specificity of the δ-endotoxins is the result of the efficiency of the various steps involved in producing an active toxin protein and its subsequent interaction with the epithelial cells in the insect mid-gut. To be insecticidal, most known δ-endotoxins must first be ingested by the insect and proteolytically activated to form an active toxin. Activation of the insecticidal crystal proteins is a multi-step process. After ingestion, the crystals must first be solubilized in the insect gut. Once solubilized, the δ-endotoxins are activated by specific proteolytic cleavages. The proteases in the insect gut can play a role in specificity by determining where the δ-endotoxin is processed. Once the δ-endotoxin has been solubilized and processed it binds to specific receptors on the surface of the insects' mid-gut epithelium and subsequently integrates into the lipid bilayer of the brush border membrane. Ion channels then form disrupting the normal function of the midgut eventually leading to the death of the insect.

In Lepidoptera, gut proteases process δ-endotoxins from 130-140 kDa protoxins to toxic proteins of approximately 60-70 kDa. Processing of the protoxin to toxin has been reported to proceed by removal of both N- and C-terminal amino acids with the exact location of processing being dependent on the specific insect gut fluids involved (Ogiwara *et al.*, 1992, J. Invert. Pathol. 60:121-126). The proteolytic activation of a δ-endotoxin can play a significant role in determining its specificity. For example, a δ-endotoxin from Bt var. *aizawa*, called IC1, has been classified as a Cry1Ab protein based on its sequence homology with other known Cry1Ab proteins. Cry1Ab proteins are typically active against lepidopteran insects. However, the IC1 protein has activity against both lepidopteran and dipteran insects depending upon how the protein is processed (Haider *et al.* 1986, Euro. J. Biochem. 156: 531-540). In a dipteran gut, a 53 kDa active IC1 toxin is obtained, whereas in a lepidopteran gut, a 55 kDa active IC1 toxin is obtained. IC1 differs from the holotype HD-1 Cry1Ab protein by only four amino acids, so gross changes in the receptor binding region do not seem to account for the differences in activity. The different proteolytic cleavages in the two different insect guts possibly allow the activated molecules to fold differently thus exposing different regions capable of binding different receptors. The specificity therefore, appears to reside with the gut proteases of the different insects.

Coleopteran insects have guts that are more neutral to acidic and coleopteran-specific δ-endotoxins are similar to the size of the activated lepidopteran-specific toxins. Therefore, the processing of coleopteran-specific δ-endotoxins was formerly considered unnecessary for toxicity. However, recent data suggests that coleopteran-active δ-endotoxins are solubilized and proteolyzed to smaller toxic polypeptides. The 73 kDa Cry3A δ-endotoxin protein produced by *B. thuringiensis* var. *tenebrionis* is readily processed in the bacterium at the N-terminus, losing 49-57 residues during or after crystal formation to produce the commonly isolated 67 kDa form (Carroll *et al.*, 1989, Biochem. J. 261:99-105). McPherson *et al.*, 1988 (Biotechnology 6:61-66) also demonstrated that the native *cry*3A gene contains two functional translational initiation codons in the same reading frame, one coding for the 73 kDa protein and the other coding for the 67 kDa protein starting at Met-1 and Met-48 respectively, of the deduced amino acid sequence (See SEQ ID NO: 2). Both proteins then can be considered naturally occurring full-length Cry3A proteins. Treatment of soluble 67 kDa Cry3A protein

with either trypsin or insect gut extract results in a cleavage product of 55 kDa with Asn-159 of the deduced amino acid sequence at the N-terminus. This polypeptide was found to be as toxic to a susceptible coleopteran insect as the native 67 kDa Cry3A toxin. (Carroll *et al. Ibid*). Thus, a natural trypsin recognition site exists between Arg-158 and Asn-159 of the deduced amino acid sequence of the native Cry3A toxin (SEQ ID NO: 2). Cry3A can also be cleaved by chymotrypsin, resulting in three polypeptides of 49, 11, and 6 kDa. N-terminal analysis of the 49 and 6 kDa components showed the first amino acid residue to be Ser-162 and Tyr-588, respectively (Carroll *et al.*, 1997 J. Invert. Biol. 70:41-49). Thus, natural chymotrypsin recognition sites exist in Cry3A between His-161 and Ser-162 and between Tyr-587 and Tyr-588 of the deduced amino acid sequence (SEQ ID NO: 2). The 49 kDa chymotrypsin product appears to be more soluble at neutral pH than the native 67 kDa protein or the 55 kDa trypsin product and retains full insecticidal activity against the Cry3A-susceptible insects, Colorado potato beetle and mustard beetle, (*Phaedon cochleariae*).

Insect gut proteases typically function in aiding the insect in obtaining needed amino acids from dietary protein. The best understood insect digestive proteases are serine proteases that appear to be the most common (Englemann and Geraerts, 1980, J. Insect Physiol. 261:703-710), particularly in lepidopteran species. The majority of coleopteran larvae and adults, for example Colorado potato beetle, have slightly acidic midguts, and cysteine proteases provide the major proteolytic activity (Wolfson and Mudock, 1990, J. Chem. Ecol. 16:1089-1102). More precisely, Thie and Houseman (1990, Insect Biochem. 20:313-318) identified and characterized the cysteine proteases, cathepsin B and H, and the aspartyl protease, cathepsin D in Colorado potato beetle. Gillikin *et al.* (1992, Arch. Insect Biochem. Physiol. 19:285-298) characterized the proteolytic activity in the guts of western corn rootworm larvae and found 15, primarily cysteine, proteases. Until disclosed in this invention, no reports have indicated that the serine protease, cathepsin G, exists in western corn rootworm. The diversity and different activity levels of the insect gut proteases may influence an insect's sensitivity to a particular Bt toxin.

Many new and novel Bt strains and δ-endotoxins with improved or novel biological activities have been described over the past five years including strains active against nematodes (EP 0517367A1). However, relatively few of these strains and toxins have activity against coleopteran insects. Further, none of the now known coleopteran-active δ-endotoxins, for

example Cry3A, Cry3B, Cry3C, Cry7A, Cry8A, Cry8B, and Cry8C, have sufficient oral toxicity against corn rootworm to provide adequate field control if delivered, for example, through microbes or transgenic plants. Therefore, other approaches for producing novel toxins active against corn rootworm need to be explored.

As more knowledge has been gained as to how the δ-endotoxins function, attempts to engineer δ-endotoxins to have new activities have increased. Engineering δ-endotoxins was made more possible by the solving of the three dimensional structure of Cry3A in 1991 (Li *et al.*, 1991, Nature 353:815-821). The protein has three structural domains: the N-terminal domain I, from residues 1-290, consists of 7 alpha helices, domain II, from residues 291-500, contains three beta-sheets and the C-terminal domain III, from residues 501-644, is a beta-sandwich. Based on this structure, a hypothesis has been formulated regarding the structure/function relationship of the δ-endotoxins. It is generally thought that domain I is primarily responsible for pore formation in the insect gut membrane (Gazit and Shai, 1993, Appl. Environ. Microbiol. 57:2816-2820), domain II is primarily responsible for interaction with the gut receptor (Ge *et al.*, 1991, J. Biol. Chem. 32:3429-3436) and that domain III is most likely involved with protein stability (Li *et al.* 1991, supra) as well as having a regulatory impact on ion channel activity (Chen *et al.*, 1993, PNAS 90:9041-9045).

Lepidopteran-active δ-endotoxins have been engineered in attempts to improve specific activity or to broaden the spectrum of insecticidal activity. For example, the silk moth (*Bombyx mori*) specificity domain from Cry1Aa was moved to Cry1Ac, thus imparting a new insecticidal activity to the resulting chimeric protein (Ge *et al.* 1989, PNAS 86: 4037-4041). Also, Bosch *et al.* 1998 (US Patent 5,736,131), created a new lepidopteran-active toxin by substituting domain III of Cry1E with domain III of Cry1C thus producing a Cry1E-Cry1C hybrid toxin with a broader spectrum of lepidopteran activity.

Several attempts at engineering the coleopteran-active δ-endotoxins have been reported. Van Rie *et al.*, 1997, (US Patent No. 5,659,123) engineered Cry3A by randomly replacing amino acids, thought to be important in solvent accessibility, in domain II with the amino acid alanine. Several of these random replacements confined to receptor binding domain II were reportedly involved in increased western corn rootworm toxicity. However, others have shown that some alanine replacements in domain II of Cry3A result in disruption of receptor binding or structural instability (Wu and Dean, 1996, J. Mol. Biol. 255: 628-640). English *et al.*, 1999,

(Intl. Pat. Appl. Publ. No. WO 99/31248) reported amino acid substitutions in Cry3Bb that caused increases in toxicity to southern and western corn rootworm. However, of the 35 reported Cry3Bb mutants, only three, with mutations primarily in domain II and the domain II-domain I interface, were active against western corn rootworm. Further, the differences in toxicity of wild-type Cry3Bb against western corn rootworm in the same assays were greater than any of the differences between the mutated Cry3Bb toxins and the wild-type Cry3Bb. Therefore, improvements in toxicity of the Cry3Bb mutants appear to be confined primarily to southern corn rootworm.

There remains a need to design new and effective pest control agents that provide an economic benefit to farmers and that are environmentally acceptable. Particularly needed are modified Cry3A toxins that control western corn rootworm, the major pest of corn in the United States, that are or could become resistant to existing insect control agents. Furthermore, agents whose application minimizes the burden on the environment, as through transgenic plants, are desirable.

In view of these needs, it is an object of the present invention to provide novel nucleic acid sequences encoding modified Cry3A toxins having increased toxicity to corn rootworm. By inserting a protease recognition site that is recognized by a target-insect gut protease in at least one position of a Cry3A toxin, in accordance with the present invention, a modified Cry3A toxin having significantly greater toxicity, particularly to western and northern corn rootworm is designed. The invention is further drawn to the novel modified Cry3A toxins resulting from the expression of the nucleic acid sequences, and to compositions and formulations containing the modified Cry3A toxins, which are capable of inhibiting the ability of insect pests to survive, grow and reproduce, or of limiting insect-related damage or loss to crop plants. The invention is further drawn to a method of making the modified Cry3A toxins and to methods of using the modified *cry*3A nucleic acid sequences, for example in microorganisms to control insects or in transgenic plants to confer protection from insect damage, and to a method of using the modified Cry3A toxins, and compositions and formulations comprising the modified Cry3A toxins, for example applying the modified Cry3A toxins or compositions or formulations to insect-infested areas, or to prophylactically treat insect-susceptible areas or plants to confer protection against the insect pests.

The novel modified Cry3A toxins described herein are highly active against insects. For example, the modified Cry3A toxins of the present invention can be used to control economically important insect pests such as western corn rootworm (*Diabrotica virgifera virgifera*) and northern corn rootworm (*D. longicornis barberi*). The modified Cry3A toxins can be used singly or in combination with other insect control strategies to confer maximal pest control efficiency with minimal environmental impact.

According to one aspect, the present invention provides an isolated nucleic acid molecule comprising a nucleotide sequence that encodes a modified Cry3A toxin, wherein the modified Cry3A toxin comprises at least one additional protease recognition site that does not naturally occur in a Cry3A toxin. The additional protease recognition site, which is recognized by a gut protease of a target insect, is inserted at approximately the same position as a naturally occurring protease recognition site in the Cry3A toxin. The modified Cry3A toxin causes higher mortality to a target insect than the mortality caused by a Cry3A toxin to the same target insect. Preferably, the modified Cry3A toxin causes at least about 50 % mortality to a target insect to which a Cry3A toxin causes only up to about 30% mortality.

In one embodiment of this aspect, the gut protease of a target insect is selected from the group consisting of serine proteases, cysteine proteases and aspartic proteases. Preferable serine proteases according to this embodiment include cathepsin G, trypsin, chymotrypsin, carboxypeptidase, endopeptidase and elastase, most preferably cathepsin G.

In another embodiment of this aspect, the additional protease recognition site is inserted in either domain I or domain III or in both domain I and domain III of the Cry3A toxin. Preferably, the additional protease recognition site is inserted in either domain I or domain III or in both domain I and domain III at a position that replaces, is adjacent to, or is within a naturally occurring protease recognition site.

In a yet another embodiment, the additional protease recognition site is inserted in domain I between amino acids corresponding to amino acid numbers 154 and 162 of SEQ ID NO: 2. Preferably, the additional protease recognition site is inserted between amino acid numbers 154 and 162 of SEQ ID NO: 2 or between amino acid numbers 107 and 115 of SEQ ID NO: 4.

In still another embodiment, the additional protease recognition site is inserted between amino acids corresponding to amino acid numbers 154 and 160 of SEQ ID NO: 2. Preferably, the

additional protease recognition site is inserted between amino acid numbers 154 and 160 of SEQ ID NO: 2 or between amino acid numbers 107 and 113 of SEQ ID NO: 4.

In a further embodiment, the additional protease recognition site is inserted in domain I between amino acids corresponding to amino acid numbers 154 and 158 of SEQ ID NO: 2. Preferably, the additional protease recognition site is inserted in domain I between amino acid numbers 154 and 158 of SEQ ID NO: 2 or between amino acid numbers 107 and 111 of SEQ ID NO: 4.

In another embodiment, the additional protease recognition site is inserted in domain III between amino acids corresponding to amino acid numbers 583 and 589 of SEQ ID NO: 2. Preferably, the additional protease site is inserted in domain III between amino acid numbers 583 and 589 of SEQ ID NO: 2 or between amino acid numbers 536 and 542 of SEQ ID NO: 4.

In still another embodiment, the additional protease recognition site is inserted in domain III between amino acids corresponding to amino acid numbers 583 and 588 of SEQ ID NO: 2. Preferably, the additional protease site is inserted in domain III between amino acid numbers 583 and 588 of SEQ ID NO: 2 or between amino acid numbers 536 and 541of SEQ ID NO: 4.

In yet another embodiment, the additional protease recognition site is inserted in domain III between amino acids corresponding to amino acid numbers 587 and 588 of SEQ ID NO: 2. Preferably, the additional protease site is inserted in domain III between amino acid numbers 587 and 588 of SEQ ID NO: 2 or between amino acid numbers 540 and 541 of SEQ ID NO: 4.

In one embodiment, the additional protease recognition site is inserted in domain I and domain III of the unmodified Cry3A toxin. Preferably, the additional protease recognition site is inserted in domain I at a position that replaces or is adjacent to a naturally occurring protease recognition site and in domain III at a position that is within, replaces, or is adjacent to a naturally occurring protease recognition site.

In another embodiment, the additional protease recognition site is inserted in domain I between amino acids corresponding to amino acid numbers 154 and 160 and in domain III between amino acids corresponding to amino acid numbers 587 and 588 of SEQ ID NO: 2. Preferably, the additional protease recognition site is inserted in domain I between amino acid numbers 154 and 160 and in domain III between amino acid numbers 587 and 588 of SEQ ID

NO: 2 or in domain I between amino acid numbers 107 and 113 and in domain III between amino acid numbers 540 and 541 of SEQ ID NO: 4.

In yet another embodiment, the additional protease recognition site is located in domain I between amino acids corresponding to amino acid numbers 154 and 158 and in domain III between amino acids corresponding to amino acid numbers 587 and 588 of SEQ ID NO: 2. Preferably, the additional protease recognition site is inserted in domain I between amino acid numbers 154 and 158 and in domain III between amino acid numbers 587 and 588 of SEQ ID NO: 2 or in domain I between amino acid numbers 107 and 111 and in domain III between amino acid numbers 540 and 541 of SEQ ID NO: 4.

In another embodiment, the additional protease recognition site is located in domain I between amino acids corresponding to amino acid numbers 154 and 158 and in domain III between amino acids corresponding to amino acid numbers 583 and 588 of SEQ ID NO: 2. Preferably, the additional protease recognition site is inserted in domain I between amino acid numbers 154 and 158 and in domain III between amino acid numbers 583 and 588 of SEQ ID NO: 2 or in domain I between amino acid numbers 107 and 111 and in domain III between amino acid numbers 536 and 541 of SEQ ID NO: 4.

In a preferred embodiment, the isolated nucleic acid molecule of the present invention comprises nucleotides 1-1791of SEQ ID NO: 6, nucleotides 1-1806 of SEQ ID NO: 8, nucleotides 1-1818 of SEQ ID NO: 10, nucleotides 1-1794 of SEQ ID NO: 12, nucleotides 1-1812 of SEQ ID NO: 14, nucleotides 1-1812 of SEQ ID NO: 16, nucleotides 1-1818 of SEQ ID NO: 18, or nucleotides 1-1791 of SEQ ID NO: 20.

In another preferred embodiment, the isolated nucleic acid molecule of the invention encodes a modified Cry3A toxin comprising the amino acid sequence set forth in SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, or SEQ ID NO: 21.

According to one embodiment of the invention, the isolated nucleic acid molecule encodes a modified Cry3A toxin that is active against a coleopteran insect. Preferably, the modified Cry3A toxin has activity against western corn rootworm.

The present invention provides a chimeric gene comprising a heterologous promoter sequence operatively linked to the nucleic acid molecule of the invention. The present invention also provides a recombinant vector comprising such a chimeric gene. Further, the present invention

provides a transgenic non-human host cell comprising such a chimeric gene. A transgenic host cell according to this aspect of the invention may be a bacterial cell or a plant cell, preferably, a plant cell. The present invention further provides a transgenic plant comprising such a plant cell. A transgenic plant according to this aspect of the invention may be sorghum, wheat, sunflower, tomato, potato, cole crops, cotton, rice, soybean, sugar beet, sugarcane, tobacco, barley, oilseed rape, or maize, preferably, maize. The present invention also provides seed from the group of transgenic plants consisting of sorghum, wheat, sunflower, tomato, potato, cole crops, cotton, rice, soybean, sugar beet, sugarcane, tobacco, barley, oilseed rape, and maize. In a particularly preferred embodiment, the seed is from a transgenic maize plant.

In another aspect, the present invention provides toxins produced by the expression of the nucleic acid molecules of the present invention. In a preferred embodiment, the toxin is produced by the expression of the nucleic acid molecule comprising nucleotides 1-1791of SEQ ID NO: 6, nucleotides 1-1806 of SEQ ID NO: 8, nucleotides 1-1818 of SEQ ID NO: 10, nucleotides 1-1794 of SEQ ID NO: 12, nucleotides 1-1812 of SEQ ID NO: 14, nucleotides 1-1812 of SEQ ID NO: 16, nucleotides 1-1818 of SEQ ID NO: 18, or nucleotides 1-1791 of SEQ ID NO: 20.

In another embodiment, the toxins of the invention are active against coleopteran insects, preferably against western corn rootworm.

In one embodiment, a toxin of the present invention comprises the amino acid sequence set forth in SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, or SEQ ID NO: 21.

The present invention also provides a composition comprising an effective insect-controlling amount of a toxin according to the invention.

In another aspect, the present invention provides a method of producing a toxin that is active against insects, comprising: (a) obtaining a host cell comprising a chimeric gene, which itself comprises a heterologous promoter sequence operatively linked to the nucleic acid molecule of the invention; and (b) expressing the nucleic acid molecule in the transgenic host cell, which results in at least one toxin that is active against insects.

In a further aspect, the present invention provides a method of producing an insect-resistant transgenic plant, comprising introducing a nucleic acid molecule of the invention into the transgenic plant, wherein the nucleic acid molecule is expressible in the transgenic plant in an

effective amount to control insects. In a preferred embodiment, the insects are coleopteran insects, preferably western corn rootworm.

In yet a further aspect, the present invention provides a method of controlling insects, comprising delivering to the insects an effective amount of a toxin of the invention. According to one embodiment, the insects are coleopteran insects, preferably, western corn rootworm. Preferably, the toxin is delivered to the insects orally. In one preferred embodiment, the toxin is delivered orally through a transgenic plant comprising a nucleic acid sequence that expresses a toxin of the present invention.

Also provided by the present invention is a method of making a modified Cry3A toxin, comprising: (a) obtaining a *cry*3A toxin gene which encodes a Cry3A toxin; (b) identifying a gut protease of a target insect; (c) obtaining a nucleotide sequence which encodes a recognition sequence for the gut protease; (d) inserting the nucleotide sequence of (c) into either domain I or domain III or both domain I and domain III at a position that replaces, is within, or adjacent to a nucleotide sequence that codes for a naturally occurring protease recognition site in a *cry*3A toxin gene, thus creating a modified *cry*3A toxin gene; (e) inserting the modified *cry*3A toxin gene in an expression cassette; (f) expressing the modified *cry*3A toxin gene in a non-human host cell, resulting in the host cell producing a modified Cry3A toxin; and, (g) bioassaying the modified Cry3A toxin against a target insect, whereby the modified Cry3A toxin causes higher mortality to the target insect than the mortality caused by a Cry3A toxin. In a preferred embodiment, the modified Cry3A toxin causes at least about 50% mortality to the target insect when the Cry3A toxin causes up to about 30% mortality.

The present invention further provides a method of controlling insects wherein the transgenic plant further comprises a second nucleic acid sequence or groups of nucleic acid sequences that encode a second pesticidal principle. Particularly preferred second nucleic acid sequences are those that encode a δ-endotoxin, those that encode a Vegetative Insecticidal Protein toxin, disclosed in U.S. Patents 5,849,870 and 5,877,012, incorporated herein by reference, or those that encode a pathway for the production of a non-proteinaceous pesticidal principle.

Yet another aspect of the present invention is the provision of a method for mutagenizing a nucleic acid molecule according to the present invention, wherein the nucleic acid molecule has been cleaved into populations of double-stranded random fragments of a desired size, comprising: (a) adding to the population of double-stranded random fragments one or more

single- or double-stranded oligonucleotides, wherein the oligonucleotides each comprise an area of identity and an area of heterology to a double-stranded template polynucleotide; (b) denaturing the resultant mixture of double-stranded random fragments and oligonucleotides into single-stranded fragments; (c) incubating the resultant population of single-stranded fragments with polymerase under conditions which result in the annealing of the single-stranded fragments at the areas of identity to form pairs of annealed fragments, the areas of identity being sufficient for one member of the pair to prime replication of the other, thereby forming a mutagenized double-stranded polynucleotide; and (d) repeating the second and third steps for at least two further cycles, wherein the resultant mixture in the second step of a further cycle includes the mutagenized double-stranded polynucleotide from the third step of the previous cycle, and wherein the further cycle forms a further mutagenized double-stranded polynucleotide.

Other aspects and advantages of the present invention will become apparent to those skilled in the art from a study of the following description of the invention and non-limiting examples.

SEQ ID NO: 1 is the native *cry*3A coding region.

SEQ ID NO: 2 is the amino acid sequence of the Cry3A toxin encoded by the native *cry*3A gene.

SEQ ID NO: 3 is the maize optimized *cry*3A coding region beginning at nucleotide 144 of the native *cry*3A coding region.

SEQ ID NO: 4 is the amino acid sequence of the Cry3A toxin encoded by the maize optimized *cry*3A gene.

SEQ ID NO: 5 is the nucleotide sequence of pCIB6850.

SEQ ID NO: 6 is the maize optimized modified *cry*3A054 coding sequence.

SEQ ID NO: 7 is the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 6.

SEQ ID NO: 8 is the maize optimized modified *cry*3A055 coding sequence.

SEQ ID NO: 9 is the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 8.

SEQ ID NO: 10 is the maize optimized modified *cry*3A085 coding sequence.

SEQ ID NO: 11 is the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 10.

SEQ ID NO: 12 is the maize optimized modified *cry*3A082 coding sequence.

SEQ ID NO: 13 is the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 12.

SEQ ID NO: 14 is the maize optimized modified *cry*3A058 coding sequence.

SEQ ID NO: 15 is the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 14.

SEQ ID NO: 16 is the maize optimized modified *cry*3A057 coding sequence.

SEQ ID NO: 17 is the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 16.

SEQ ID NO: 18 is the maize optimized modified *cry*3A056 coding sequence.

SEQ ID NO: 19 is the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 18.

SEQ ID NO: 20 is the maize optimized modified *cry*3A083 coding sequence.

SEQ ID NO: 21 is the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 20.

SEQ ID NOS: 22-34 are PCR primers useful in the present invention.

SEQ ID NO: 35 is an amino acid sequence comprising a cathepsin G recognition site.

SEQ ID NO: 36 is an amino acid sequence comprising a cathepsin G recognition site.

SEQ ID NO: 37 is an amino acid sequence comprising a cathepsin G recognition site.

SEQ ID NO: 38 is an amino acid sequence comprising a cathepsin G recognition site.

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

"Activity" of the modified Cry3A toxins of the invention is meant that the modified Cry3A toxins function as orally active insect control agents, have a toxic effect, or are able to disrupt or deter insect feeding, which may or may not cause death of the insect. When a modified Cry3A toxin of the invention is delivered to the insect, the result is typically death of the insect, or the insect does not feed upon the source that makes the modified Cry3A toxin available to the insect.

"Adjacent to"- According to the present invention, an additional protease recognition site is "adjacent to" a naturally occurring protease recognition site when the additional protease recognition site is within four residues, preferably within three residues, more preferably within two residues, and most preferably within one residue of a naturally occurring protease recognition site. For example, an additional protease recognition site inserted between Pro-154 and Arg-158 of the deduced amino acid sequence of a Cry3A toxin (SEQ ID NO: 2) is "adjacent to" the naturally occurring trypsin recognition site located between Arg-158 and Asn-159 of the deduced amino acid sequence of the Cry3A toxin (SEQ ID NO: 2).

The phrase "approximately the same position" as used herein to describe the location where an additional protease recognition site is inserted into a Cry3A toxin in relation to a naturally occurring protease recognition site, means that the location is at most four residues away from a naturally occurring protease recognition site. The location can also be three or two residues away from a naturally occurring protease recognition site. The location can also be one residue away from a naturally occurring protease recognition site. "Approximately the same position" can also mean that the additional protease recognition site is inserted within a naturally occurring protease recognition site.

"Associated with / operatively linked" refer to two nucleic acid sequences that are related physically or functionally. For example, a promoter or regulatory DNA sequence is said to be "associated with" a DNA sequence that codes for an RNA or a protein if the two sequences are operatively linked, or situated such that the regulatory DNA sequence will affect the expression level of the coding or structural DNA sequence.

A "chimeric gene" or "chimeric construct" is a recombinant nucleic acid sequence in which a promoter or regulatory nucleic acid sequence is operatively linked to, or associated with, a nucleic acid sequence that codes for an mRNA or which is expressed as a protein, such that the regulatory nucleic acid sequence is able to regulate transcription or expression of the associated nucleic acid coding sequence. The regulatory nucleic acid sequence of the chimeric gene is not normally operatively linked to the associated nucleic acid sequence as found in nature.

A "coding sequence" is a nucleic acid sequence that is transcribed into RNA such as mRNA, rRNA, tRNA, snRNA, sense RNA or antisense RNA. Preferably the RNA is then translated in an organism to produce a protein.

To "control" insects means to inhibit, through a toxic effect, the ability of insect pests to survive, grow, feed, and/or reproduce, or to limit insect-related damage or loss in crop plants. To "control" insects may or may not mean killing the insects, although it preferably means killing the insects.

Corresponding to: in the context of the present invention, "corresponding to" means that when the amino acid sequences of variant Cry3A δ-endotoxins are aligned with each other, the amino acids that "correspond to" certain enumerated positions in the present invention are those that align with these positions in the Cry3A toxin (SEQ ID NO: 2), but that are not necessarily in these exact numerical positions relative to the particular Cry3A amino acid sequence of the invention. For example, the maize optimized *cry*3A gene (SEQ ID NO: 3) of the invention encodes a Cry3A toxin (SEQ ID NO: 4) that begins at Met-48 of the Cry3A toxin (SEQ ID NO: 2) encoded by the native *cry*3A gene (SEQ ID NO: 1). Therefore, according to the present invention, amino acid numbers 107-115, including all numbers in between, and 536-541, including all numbers in between, of SEQ ID NO: 4 correspond to amino acid numbers 154-163, and all numbers in between, and 583-588, and all numbers in between, respectively, of SEQ ID NO: 2.

A "Cry3A toxin", as used herein, refers to an approximately 73 kDa *Bacillus thuringiensis* var. *tenebrionis* (Kreig *et al.*, 1983, Z. Angew. Entomol. 96:500-508) (Bt) coleopteran-active protein (Sekar *et al.*, 1987, Proc. Nalt. Acad. Sci. 84:7036-7040), for example SEQ ID NO: 2, as well as any truncated lower molecular weight variants, derivable from a Cry3A toxin, for example SEQ ID NO: 4, and retaining substantially the same toxicity as the Cry3A toxin. The lower molecular weight variants can be obtained by protease cleavage of naturally occurring protease recognition sites of the Cry3A toxin or by a second translational initiation codon in the same frame as the translational initiation codon coding for the 73 kDa Cry3A toxin. The amino acid sequence of a Cry3A toxin and the lower molecular weight variants thereof can be found in a toxin naturally occurring in Bt. A Cry3A toxin can be encoded by a native Bt gene as in SEQ ID NO: 1 or by a synthetic coding sequence as in SEQ ID NO: 3. A "Cry3A toxin" does not have any additional protease recognition sites over the protease recognition sites that naturally occur in the Cry3A toxin. A Cry3A toxin can be isolated, purified or expressed in a heterologous system.

A "*cry*3A gene", as used herein, refers to the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO: 3. A *cry*3A gene (Sekar *et al.*, 1987, Proc. Natl. Acad. Sci. 84:7036-7040) can be naturally occurring, as found in *Bacillus thuringiensis* var. *tenebrionis* (Krcig *et al.*, 1983, Z. Angew. Entomol. 96:500-508), or synthetic and encodes a Cry3A toxin. The *cry*3A gene of this invention can be referred to as the native *cry*3A gene as in SEQ ID NO: 1 or the maize-optimized *cry*3A gene as in SEQ ID NO: 3.

To "deliver" a toxin means that the toxin comes in contact with an insect, resulting in toxic effect and control of the insect. The toxin can be delivered in many recognized ways, e.g., orally by ingestion by the insect or by contact with the insect via transgenic plant expression, formulated protein composition(s), sprayable protein composition(s), a bait matrix, or any other art-recognized toxin delivery system.

"Effective insect-controlling amount" means that concentration of toxin that inhibits, through a toxic effect, the ability of insects to survive, grow, feed and/or reproduce, or to limit insect-related damage or loss in crop plants. "Effective insect-controlling amount" may or may not mean killing the insects, although it preferably means killing the insects.

"Expression cassette" as used herein means a nucleic acid sequence capable of directing expression of a particular nucleotide sequence in an appropriate host cell, comprising a promoter operably linked to the nucleotide sequence of interest which is operably linked to termination signals. It also typically comprises sequences required for proper translation of the nucleotide sequence. 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 that 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 nucleic acid 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 that 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.

A "gene" is a defined region that is located within a genome and that, besides the aforementioned coding nucleic acid sequence, comprises other, primarily regulatory, nucleic acid sequences responsible for the control of the expression, that is to say the transcription and translation, of the coding portion. A gene may also comprise other 5' and 3' untranslated sequences and termination sequences. Further elements that may be present are, for example, introns.

"Gene of interest" refers to any gene which, when transferred to a plant, confers upon the plant a desired characteristic such as antibiotic resistance, virus resistance, insect resistance, disease resistance, or resistance to other pests, herbicide tolerance, improved nutritional value, improved performance in an industrial process or altered reproductive capability. The "gene of interest" may also be one that is transferred to plants for the production of commercially valuable enzymes or metabolites in the plant.

A "gut protease" is a protease naturally found in the digestive tract of an insect. This protease is usually involved in the digestion of ingested proteins.

A "heterologous" nucleic acid sequence is a nucleic acid sequence not naturally associated with a host cell into which it is introduced, including non- naturally occurring multiple copies of a naturally occurring nucleic acid sequence.

A "homologous" nucleic acid sequence is a nucleic acid sequence naturally associated with a host cell into which it is introduced.

"Homologous recombination" is the reciprocal exchange of nucleic acid fragments between homologous nucleic acid molecules.

"Insecticidal" is defined as a toxic biological activity capable of controlling insects, preferably by killing them.

A nucleic acid sequence is "isocoding with" a reference nucleic acid sequence when the nucleic acid sequence encodes a polypeptide having the same amino acid sequence as the polypeptide encoded by the reference nucleic acid sequence.

An "isolated" nucleic acid molecule or an isolated toxin is a nucleic acid molecule or toxin that, by the hand of man, exists apart from its native environment and is therefore not a product of nature. An isolated nucleic acid molecule or toxin may exist in a purified form or may exist in a non-native environment such as, for example, a recombinant host cell.

A "modified Cry3A toxin" of this invention, refers to a Cry3A-derived toxin having at least one additional protease recognition site that is recognized by a gut protease of a target insect, which does not naturally occur in a Cry3A toxin. A modified Cry3A toxin is not naturally occurring and, by the hand of man, comprises an amino acid sequence that is not identical to a naturally occurring toxin found in *Bacillus thuringiensis*. The modified Cry3A toxin causes higher mortality to a target insect than the mortality caused by a Cry3A toxin to the same target insect.

A "modified *cry*3A gene" according to this invention, refers to a *cry*3A-derived gene comprising the coding sequence of at least one additional protease recognition site that does not naturally occur in an unmodified *cry*3A gene. The modified *cry*3A gene can be derived from a native *cry*3A gene or from a synthetic *cry*3A gene.

A "naturally occurring protease recognition site" is a location within a Cry3A toxin that is cleaved by a non-insect derived protease or by a protease or gut extract from an insect species susceptible to the Cry3A toxin. For example, a naturally occurring protease recognition site, recognized by trypsin and proteases found in a susceptible insect gut extract, exists between Arg-158 and Asn-159 of the deduced Cry3A toxin amino acid sequence (SEQ ID NO: 2). Naturally occurring protease recognition sites, recognized by chymotrypsin, exist between His-161 and Ser-162 as well as between Tyr-587 and Tyr-588 of the deduced Cry3A toxin amino acid sequence (SEQ ID NO: 2).

A "nucleic acid molecule" or "nucleic acid sequence" is a linear segment of single- or double-stranded DNA or RNA that can be isolated from any source. In the context of the present invention, the nucleic acid molecule is preferably a segment of DNA.

A "plant" is any plant at any stage of development, particularly a seed plant.

A "plant cell" is a structural and physiological unit of a plant, comprising a protoplast and a cell wall. The plant cell may be in the form of an isolated single cell or a cultured cell, or as a part of a higher organized unit such as, for example, plant tissue, a plant organ, or a whole plant.

"Plant cell culture" means cultures of plant units such as, for example, protoplasts, cell culture cells, cells in plant tissues, pollen, pollen tubes, ovules, embryo sacs, zygotes and embryos at various stages of development.

"Plant material" refers to leaves, stems, roots, flowers or flower parts, fruits, pollen, egg cells, zygotes, seeds, cuttings, cell or tissue cultures, or any other part or product of a plant.

A "plant organ" is a distinct and visibly structured and differentiated part of a plant such as a root, stem, leaf, flower bud, or embryo.

"Plant tissue" as used herein means a group of plant cells organized into a structural and functional unit. Any tissue of a plant *in planta* or in culture is included. This term includes, but is not limited to, whole plants, plant organs, plant seeds, tissue culture and any groups of plant cells organized into structural and/or functional units. The use of this term in conjunction with, or in the absence of, any specific type of plant tissue as listed above or otherwise embraced by this definition is not intended to be exclusive of any other type of plant tissue.

A "promoter" is an untranslated DNA sequence upstream of the coding region that contains the binding site for RNA polymerase and initiates transcription of the DNA. The promoter region may also include other elements that act as regulators of gene expression.

A "protoplast" is an isolated plant cell without a cell wall or with only parts of the cell wall.

"Regulatory elements" refer to sequences involved in controlling the expression of a nucleotide sequence. Regulatory elements comprise a promoter operably linked to the nucleotide sequence of interest and termination signals. They also typically encompass sequences required for proper translation of the nucleotide sequence.

"Replaces" a naturally occurring protease recognition site - According to the present invention, an additional protease recognition site "replaces" a naturally occurring protease recognition site when insertion of the additional protease recognition site eliminates the naturally occurring protease recognition site. For example, an additional protease recognition site inserted between Pro-154 and Pro-160 of the deduced amino acid sequence of a Cry3A toxin (SEQ ID NO: 2) which eliminates the Arg-158 and Asn-159 residues "replaces" the naturally occurring trypsin recognition site located between Arg-158 and Asn-159 of the deduced amino acid sequence of the Cry3A toxin (SEQ ID NO: 2).

"Serine proteases", describe the same group of enzymes that catalyze the hydrolysis of covalent peptidic bonds using a mechanism based on nucleophilic attack of the targeted peptidic bond by a serine. Serine proteases are sequence specific. That is, each serine protease recognizes a specific sub-sequence within a protein where enzymatic recognition occurs.

A "target insect" is an insect pest species that has little or no susceptibility to a Cry3A toxin and is identified as being a candidate for using the technology of the present invention to control. This control can be achieved through several means but most preferably through the expression of the nucleic acid molecules of the invention in transgenic plants.

A "target insect gut protease" is a protease found in the gut of a target insect whose recognition site can be inserted into a Cry3A toxin to create a modified Cry3A toxin of the invention.

"Transformation" is a process for introducing heterologous nucleic acid into a host cell or organism. In particular, "transformation" means the stable integration of a DNA molecule into the genome of an organism of interest.

"Transformed / transgenic / recombinant" refer to a host organism such as a bacterium or a plant into which a heterologous nucleic acid molecule has been introduced. The nucleic acid molecule can be stably integrated into the genome of the host or the nucleic acid molecule can also be present as an extrachromosomal molecule. Such an extrachromosomal molecule can be auto-replicating. Transformed cells, tissues, or plants are understood to encompass not only the end product of a transformation process, but also transgenic progeny thereof. A "non-transformed", "non-transgenic", or "non- recombinant" host refers to a wild-type organism, e.g., a bacterium or plant, which does not contain the heterologous nucleic acid molecule.

"Within" a naturally occurring protease recognition site – According to the present invention, an additional protease recognition site is "within" a naturally occurring protease recognition site when the additional protease recognition site lies between the amino acid residue that comes before and the amino acid residue that comes after the naturally occurring protease recognition site. For example, an additional protease recognition site inserted between Tyr-587 and Tyr-588 of the deduced amino acid sequence of a Cry3A toxin (SEQ ID NO: 2) is "within" a naturally occurring chymotrypsin recognition site located between Tyr-587 and Tyr-588 of the deduced amino acid sequence of the Cry3A toxin (SEQ ID NO: 2). The insertion of an additional protease recognition site within a naturally occurring protease recognition site may or may not change the recognition of the naturally occurring protease recognition site by a protease.

Nucleotides are indicated by their bases by the following standard abbreviations: adenine (A), cytosine (C), thymine (T), and guanine (G). Amino acids are likewise indicated by the

following standard abbreviations: alanine (Ala; A), arginine (Arg; R), asparagine (Asn; N), aspartic acid (Asp; D), cysteine (Cys; C), glutamine (Gln; Q), glutamic acid (Glu; E), glycine (Gly; G), histidine (His; H), isoleucine (Ile; I), leucine (Leu; L), lysine (Lys; K), methionine (Met; M), phenylalanine (Phe; F), proline (Pro; P), serine (Ser; S), threonine (Thr; T), tryptophan (Trp; W), tyrosine (Tyr; Y), and valine (Val; V).

This invention relates to modified *cry*3A nucleic acid sequences whose expression results in modified Cry3A toxins, and to the making and using of the modified Cry3A toxins to control insect pests. The expression of the modified *cry*3A nucleic acid sequences results in modified Cry3A toxins that can be used to control coleopteran insects such as western corn rootworm and northern corn rootworm. A modified Cry3A toxin of the present invention comprises at least one additional protease recognition site that does not naturally occur in a Cry3A toxin. The additional protease recognition site, which is recognized by a gut protease of a target insect, is inserted at approximately the same position as a naturally occurring protease recognition site in a Cry3A toxin. The modified Cry3A toxin causes higher mortality to a target insect than the mortality caused by a Cry3A toxin to the same target insect. Preferably, the modified Cry3A toxin causes at least about 50 % mortality to the target insect to which a Cry3A toxin causes up to about 30% mortality.

In one preferred embodiment, the invention encompasses an isolated nucleic acid molecule that encodes a modified Cry3A toxin, wherein the additional protease recognition site is recognized by the target insect gut protease, cathepsin G. Cathepsin G activity is determined to be present in the gut of the target insect, western corn rootworm, as described in Example 2. Preferably, the substrate amino acid sequence, AAPF (SEQ ID NO: 35), used to determine the presence of the cathepsin G activity is inserted into the Cry3A toxin according to the present invention. Other cathepsin G recognition sites can also be used according to the present invention, for example, AAPM (SEQ ID NO: 36), AVPF (SEQ ID NO: 37), PFLF (SEQ ID NO: 38) or other cathepsin G recognition sites as determined by the method of Tanaka *et al.*, 1985 (Biochemistry 24:2040-2047), incorporated herein by reference. Protease recognition sites of other proteases identified in a target insect gut can be used, for example, protease recognition sites recognized by other serine proteases, cysteine proteases and aspartic

proteases. Preferable serine proteases encompassed by this embodiment include trypsin, chymotrypsin, carboxypeptidase, endopeptidase and elastase.

In another preferred embodiment, the invention encompasses an isolated nucleic acid molecule that encodes a modified Cry3A toxin wherein the additional protease recognition site is inserted in either domain I or domain III or in both domain I and domain III of the Cry3A toxin. Preferably, the additional protease recognition site is inserted in domain I, domain III, or domain I and domain III at a position that replaces, is adjacent to, or is within a naturally occurring protease recognition site in the Cry3A toxin. Specifically exemplified herein are nucleic acid molecules that encode modified Cry3A toxins that comprise a cathepsin G recognition site inserted in domain I, domain III, or domain I and domain III at a position that replaces, is adjacent to, or is within a naturally occurring protease recognition site in the unmodified Cry3A toxin.

Specifically exemplified teachings of methods to make modified *cry*3A nucleic acid molecules that encode modified Cry3A toxins can be found in Example 3. Those skilled in the art will recognize that other methods known in the art can also be used to insert additional protease recognition sites into Cry3A toxins according to the present invention.

In another preferred embodiment, the invention encompasses an isolated nucleic acid molecule that encodes a modified Cry3A toxin wherein the additional protease recognition site is inserted in domain I between amino acids corresponding to amino acid numbers 154 and 162 of SEQ ID NO: 2. Preferably, the additional protease recognition site is inserted between amino acid numbers 154 and 162 of SEQ ID NO: 2 or between amino acid numbers 107 and 115 of SEQ ID NO: 4. In a preferred embodiment, the additional protease recognition site is inserted between amino acids corresponding to amino acid numbers 154 and 160 of SEQ ID NO: 2. Preferably, the additional protease recognition site is inserted between amino acid number 154 and 160 of SEQ ID NO: 2 or between amino acid numbers 107 and 113 of SEQ ID NO: 4. Specifically exemplified herein is a nucleic acid molecule, designated *cry*3A054 (SEQ ID NO: 6), that encodes the modified Cry3A054 toxin (SEQ ID NO: 7) comprising a cathepsin G recognition site inserted in domain I between amino acid numbers 107 and 113 of SEQ ID NO: 4. The cathepsin G recognition site replaces a naturally occurring trypsin recognition site and is adjacent to a naturally occurring chymotrypsin recognition site. When expressed in a heterologous host, the nucleic acid molecule of SEQ ID

NO: 6 results in insect control activity against western corn rootworm and northern corn rootworm, showing that the nucleic acid sequence set forth in SEQ ID NO: 6 is sufficient for such insect control activity.

In another preferred embodiment, the additional protease recognition site is inserted in domain I between amino acids corresponding to amino acid numbers 154 and 158 of SEQ ID NO: 2. Preferably, the additional protease recognition site is inserted in domain I between amino acid numbers 154 and 158 of SEQ ID NO: 2 or between amino acid numbers 107 and 111 of SEQ ID NO: 4. Specifically exemplified herein are nucleic acid molecules, designated *cry*3A055 (SEQ ID NO: 8), that encodes the modified Cry3A055 toxin (SEQ ID NO: 9), and *cry*3A085 (SEQ ID NO: 10), that encodes the modified Cry3A085 toxin (SEQ ID NO: 11), comprising a cathepsin G recognition site inserted in domain I between amino acid numbers 107 and 111 of SEQ ID NO: 4. The cathepsin G recognition site is adjacent to naturally occurring trypsin and chymotrypsin recognition sites. When expressed in a heterologous host, the nucleic acid molecule of SEQ ID NO: 8 or SEQ ID NO: 10 results in insect control activity against western corn rootworm and northern corn rootworm, showing that the nucleic acid sequence set forth in SEQ ID NO: 8 or SEQ ID NO: 10 is sufficient for such insect control activity.

In a preferred embodiment, the invention encompasses an isolated nucleic acid molecule that encodes a modified Cry3A toxin wherein the additional protease recognition site is inserted in domain III between amino acids corresponding to amino acid numbers 583 and 589 of SEQ ID NO: 2. Preferably, the additional protease site is inserted in domain III between amino acid numbers 583 and 589 of SEQ ID NO: 2 or between amino acid numbers 536 and 542 of SEQ ID NO: 4.

In another preferred embodiment, the invention encompasses an isolated nucleic acid molecule that encodes a modified Cry3A toxin wherein the additional protease recognition site is inserted in domain III between amino acids corresponding to amino acid numbers 583 and 588 of SEQ ID NO: 2. Preferably, the additional protease site is inserted in domain III between amino acid numbers 583 and 588 of SEQ ID NO: 2 or between amino acid numbers 536 and 541 of SEQ ID NO: 4. Specifically exemplified herein is a nucleic acid molecule, designated *cry*3A082 (SEQ ID NO: 12), that encodes the modified Cry3A082 toxin (SEQ ID NO: 13) comprising a cathepsin G recognition site inserted in domain III between amino acid numbers 536 and 541 of SEQ ID NO: 4. The cathepsin G recognition site replaces a naturally

occurring chymotrypsin recognition site. When expressed in a heterologous host, the nucleic acid molecule of SEQ ID NO: 12 results in insect control activity against western corn rootworm and northern corn rootworm, showing that the nucleic acid sequence set forth in SEQ ID NO: 12 is sufficient for such insect control activity.

In another preferred embodiment, the additional protease recognition site is inserted in domain III between amino acids corresponding to amino acid numbers 587 and 588 of SEQ ID NO: 2. Preferably, the additional protease site is inserted in domain III between amino acid numbers 587 and 588 of SEQ ID NO: 2 or between amino acid numbers 540 and 541 of SEQ ID NO: 4. Specifically exemplified herein is a nucleic acid molecule, designated *cry*3A058 (SEQ ID NO: 14), that encodes the modified Cry3A058 toxin (SEQ ID NO: 15) comprising a cathepsin G recognition site inserted in domain III between amino acid numbers 540 and 541 of SEQ ID NO: 4. The cathepsin G recognition site is within a naturally occurring chymotrypsin recognition site. When expressed in a heterologous host, the nucleic acid molecule of SEQ ID NO: 14 results in insect control activity against western corn rootworm and northern corn rootworm, showing that the nucleic acid sequence set forth in SEQ ID NO: 14 is sufficient for such insect control activity.

In yet another preferred embodiment, the invention encompasses an isolated nucleic acid molecule that encodes a modified Cry3A toxin wherein the additional protease recognition site is inserted in domain I between amino acids corresponding to amino acid numbers 154 and 160 and in domain III between amino acids corresponding to amino acid numbers 587 and 588 of SEQ ID NO: 2. Preferably, the additional protease recognition site is inserted in domain I between amino acid numbers 154 and 160 and in domain III between amino acid numbers 587 and 588 of SEQ ID NO: 2 or in domain I between amino acid numbers 107 and 113 and in domain III between amino acid numbers 540 and 541 of SEQ ID NO: 4. Specifically exemplified herein is a nucleic acid molecule, designated *cry*3A057 (SEQ ID NO: 16), that encodes the modified Cry3A057 toxin (SEQ ID NO: 17) comprising a cathepsin G recognition site inserted in domain I between amino acid numbers 107 and 113 and in domain III between amino acid numbers 540 and 541 of SEQ ID NO: 4. The cathepsin G recognition site replaces a naturally occurring trypsin recognition site and is adjacent to a naturally occurring chymotrypsin recognition site in domain I and is within a naturally occurring chymotrypsin recognition site in domain III. When expressed in a heterologous host, the

nucleic acid molecule of SEQ ID NO: 16 results in insect control activity against western corn rootworm and northern corn rootworm, showing that the nucleic acid sequence set forth in SEQ ID NO: 16 is sufficient for such insect control activity.

In yet another preferred embodiment, the additional protease recognition site is located in domain I between amino acids corresponding to amino acid numbers 154 and 158 and in domain III between amino acids corresponding to amino acid numbers 587 and 588 of SEQ ID NO: 2. Preferably, the additional protease recognition site is inserted in domain I between amino acid numbers 154 and 158 and in domain III between amino acid numbers 587 and 588 of SEQ ID NO: 2 or in domain I between amino acid numbers 107 and 111 and in domain III between amino acid numbers 540 and 541 of SEQ ID NO: 4. Specifically exemplified herein is the nucleic acid molecule designated *cry*3A056 (SEQ ID NO: 18), which encodes the modified Cry3A056 toxin (SEQ ID NO: 19) comprising a cathepsin G recognition site inserted in domain I between amino acid numbers 107 and 111 and in domain III between amino acid numbers 540 and 541 of SEQ ID NO: 4. The cathepsin G recognition site is adjacent to naturally occurring trypsin and chymotrypsin recognition sites in domain I and is within a naturally occurring chymotrypsin recognition site in domain III. When expressed in a heterologous host, the nucleic acid molecule of SEQ ID NO: 18 results in insect control activity against western corn rootworm and northern corn rootworm, showing that the nucleic acid sequence set forth in SEQ ID NO: 18 is sufficient for such insect control activity.

In still another preferred embodiment, the additional protease recognition site is located in domain I between amino acids corresponding to amino acid numbers 154 and 158 and in domain III between amino acids corresponding to amino acid numbers 583 and 588 of SEQ ID NO: 2. Preferably, the additional protease recognition site is inserted in domain I between amino acid numbers 154 and 158 and in domain III between amino acid numbers 583 and 588 of SEQ ID NO: 2 or in domain I between amino acid numbers 107 and 111 and in domain III between amino acid numbers 536 and 541 of SEQ ID NO: 4. Specifically exemplified herein is a nucleic acid molecule, designated *cry*3A083 (SEQ ID NO: 20), which encodes the modified Cry3A083 toxin (SEQ ID NO: 21) comprising a cathepsin G recognition site inserted in domain I between amino acid numbers 107 and 111 and in domain III between amino acid numbers 536 and 541 of SEQ ID NO: 4. The cathepsin G recognition site is adjacent to naturally occurring trypsin and chymotrypsin recognition sites in domain I and

replaces a naturally occurring chymotrypsin recognition site in domain III. When expressed in a heterologous host, the nucleic acid molecule of SEQ ID NO: 20 results in insect control activity against western corn rootworm and northern corn rootworm, showing that the nucleic acid sequence set forth in SEQ ID NO: 20 is sufficient for such insect control activity.

In a preferred embodiment, the isolated nucleic acid molecule of the present invention comprises nucleotides 1-1791of SEQ ID NO: 6, nucleotides 1-1806 of SEQ ID NO: 8, nucleotides 1-1812 of SEQ ID NO: 10, nucleotides 1-1794 of SEQ ID NO: 12, nucleotides 1-1818 of SEQ ID NO: 14, nucleotides 1-1812 of SEQ ID NO: 16, nucleotides 1-1791 of SEQ ID NO: 18, and nucleotides 1-1818 of SEQ ID NO: 20.

In another preferred embodiment, the invention encompasses the isolated nucleic acid molecule that encodes a modified Cry3A toxin comprising the amino acid sequence set forth in SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, or SEQ ID NO: 21.

The present invention also encompasses recombinant vectors comprising the nucleic acid sequences of this invention. In such vectors, the nucleic acid sequences are preferably comprised in expression cassettes comprising regulatory elements for expression of the nucleotide sequences in a host cell capable of expressing the nucleotides sequences. Such regulatory elements usually comprise promoter and termination signals and preferably also comprise elements allowing efficient translation of polypeptides encoded by the nucleic acid sequences of the present invention. Vectors comprising the nucleic acid sequences are usually capable of replication in particular host cells, preferably as extrachromosomal molecules, and are therefore used to amplify the nucleic acid sequences of this invention in the host cells. In one embodiment, host cells for such vectors are microorganisms, such as bacteria, in particular *Bacillus thuringiensis* or *E. coli*. In another embodiment, host cells for such recombinant vectors are endophytes or epiphytes. A preferred host cell for such vectors is a eukaryotic cell, such as a plant cell. Plant cells such as maize cells are most preferred host cells. In another preferred embodiment, such vectors are viral vectors and are used for replication of the nucleotide sequences in particular host cells, e.g. insect cells or plant cells. Recombinant vectors are also used for transformation of the nucleotide sequences of this invention into host cells, whereby the nucleotide sequences are stably integrated into the DNA of such host cells. In one, such host cells are prokaryotic cells. In a preferred embodiment,

such host cells are eukaryotic cells, such as plant cells. In a most preferred embodiment, the host cells are plant cells, such as maize cells.

In another aspect, the present invention encompasses modified Cry3A toxins produced by the expression of the nucleic acid molecules of the present invention.

In preferred embodiments, the modified Cry3A toxins of the invention comprise a polypeptide encoded by a nucleotide sequence of the invention. In a further preferred embodiment, the modified Cry3A toxin is produced by the expression of the nucleic acid molecule comprising nucleotides 1-1791of SEQ ID NO: 6, nucleotides 1-1806 of SEQ ID NO: 8, nucleotides 1-1812 of SEQ ID NO: 10, nucleotides 1-1794 of SEQ ID NO: 12, nucleotides 1-1818 of SEQ ID NO: 14, nucleotides 1-1812 of SEQ ID NO: 16, nucleotides 1-1791 of SEQ ID NO: 18, and nucleotides 1-1818 of SEQ ID NO: 20.

In a preferred embodiment, a modified Cry3A toxin of the present invention comprises the amino acid sequence set forth in SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, or SEQ ID NO: 21.

The modified Cry3A toxins of the present invention have insect control activity when tested against insect pests in bioassays. In another preferred embodiment, the modified Cry3A toxins of the invention are active against coleopteran insects, preferably against western corn rootworm and northern corn rootworm. The insect controlling properties of the modified Cry3A toxins of the invention are further illustrated in Examples 4 and 6.

The present invention also encompasses a composition comprising an effective insect-controlling amount of a modified Cry3A toxin according to the invention.

In another preferred embodiment, the invention encompasses a method of producing a modified Cry3A toxin that is active against insects, comprising: (a) obtaining a host cell comprising a chimeric gene, which itself comprises a heterologous promoter sequence operatively linked to the nucleic acid molecule of the invention; and (b) expressing the nucleic acid molecule in the transgenic host cell, which results in at least one modified Cry3A toxin that is active against insects.

In a further preferred embodiment, the invention encompasses a method of producing an insect-resistant transgenic plant, comprising introducing a nucleic acid molecule of the invention into the transgenic plant, wherein the nucleic acid molecule is expressible in the transgenic plant in an effective amount to control insects. In a preferred embodiment, the

insects are coleopteran insects, preferably western corn rootworm and northern corn rootworm.

In yet a further preferred embodiment, the invention encompasses a method of controlling insects, comprising delivering to the insects an effective amount of a modified Cry3A toxin of the invention. According to this embodiment, the insects are coleopteran insects, preferably, western corn rootworm and northern corn rootworm. Preferably, the modified Cry3A toxin is delivered to the insects orally. In one preferred aspect, the toxin is delivered orally through a transgenic plant comprising a nucleic acid sequence that expresses a modified Cry3A toxin of the present invention.

The present invention also encompasses a method of making a modified Cry3A toxin, comprising: (a) obtaining a *cry*3A toxin gene which encodes a Cry3A toxin; (b) identifying a gut protease of a target insect; (c) obtaining a nucleotide sequence which encodes a recognition site for the gut protease; (d) inserting the nucleotide sequence of (c) into either domain I or domain III or both domain I and domain III at a position that replaces, is within, or adjacent to a nucleotide sequence that codes for a naturally occurring protease recognition site in the *cry*3A toxin gene, thus creating a modified *cry*3A toxin gene; (e) inserting the modified *cry*3A toxin gene in an expression cassette; (f) expressing the modified *cry*3A toxin gene in a non-human host cell, resulting in the host cell producing a modified Cry3A toxin; and, (g) bioassaying the modified Cry3A toxin against a target insect, which causes higher mortality to the target insect than the mortality caused by a Cry3A toxin. In a preferred embodiment, the modified Cry3A toxin causes at least about 50% mortality to the target insect when the Cry3A toxin causes up to about 30% mortality.

The present invention further encompasses a method of controlling insects wherein the transgenic plant further comprises a second nucleic acid sequence or groups of nucleic acid sequences that encode a second pesticidal principle. Particularly preferred second nucleic acid sequences are those that encode a δ-endotoxin, those that encode a Vegetative Insecticidal Protein toxin, disclosed in U.S. Patents 5,849,870 and 5,877,012, incorporated herein by reference, or those that encode a pathway for the production of a non-proteinaceous principle.

In further embodiments, the nucleotide sequences of the invention can be further modified by incorporation of random mutations in a technique known as *in vitro* recombination or DNA shuffling. This technique is described in Stemmer *et al.*, Nature 370:389-391 (1994) and U.S.

Patent 5,605,793, which are incorporated herein by reference. Millions of mutant copies of a nucleotide sequence are produced based on an original nucleotide sequence of this invention and variants with improved properties, such as increased insecticidal activity, enhanced stability, or different specificity or ranges of target-insect pests are recovered. The method encompasses forming a mutagenized double-stranded polynucleotide from a template double-stranded polynucleotide comprising a nucleotide sequence of this invention, wherein the template double-stranded polynucleotide has been cleaved into double-stranded-random fragments of a desired size, and comprises the steps of adding to the resultant population of double-stranded random fragments one or more single or double-stranded oligonucleotides, wherein said oligonucleotides comprise an area of identity and an area of heterology to the double-stranded template polynucleotide; denaturing the resultant mixture of double-stranded random fragments and oligonucleotides into single-stranded fragments; incubating the resultant population of single-stranded fragments with a polymerase under conditions which result in the annealing of said single- stranded fragments at said areas of identity to form pairs of annealed fragments, said areas of identity being sufficient for one member of a pair to prime replication of the other, thereby forming a mutagenized double-stranded polynucleotide; and repeating the second and third steps for at least two further cycles, wherein the resultant mixture in the second step of a further cycle includes the mutagenized double-stranded polynucleotide from the third step of the previous cycle, and the further cycle forms a further mutagenized double-stranded polynucleotide. In a preferred embodiment, the concentration of a single species of double- stranded random fragment in the population of double-stranded random fragments is less than 1% by weight of the total DNA. In a further preferred embodiment, the template double-stranded polynucleotide comprises at least about 100 species of polynucleotides. In another preferred embodiment, the size of the double-stranded random fragments is from about 5 bp to 5 kb. In a further preferred embodiment, the fourth step of the method comprises repeating the second and the third steps for at least 10 cycles.

Expression of the Nucleotide Sequences in Heterologous Microbial Hosts

As biological insect control agents, the insecticidal modified Cry3A toxins are produced by expression of the nucleotide sequences in heterologous host cells capable of expressing the

nucleotide sequences. In a first embodiment, *B. thuringiensis* cells comprising modifications of a nucleotide sequence of this invention are made. Such modifications encompass mutations or deletions of existing regulatory elements, thus leading to altered expression of the nucleotide sequence, or the incorporation of new regulatory elements controlling the expression of the nucleotide sequence. In another embodiment, additional copies of one or more of the nucleotide sequences are added to *Bacillus thuringiensis* cells either by insertion into the chromosome or by introduction of extrachromosomally replicating molecules containing the nucleotide sequences.

In another embodiment, at least one of the nucleotide sequences of the invention is inserted into an appropriate expression cassette, comprising a promoter and termination signal. Expression of the nucleotide sequence is constitutive, or an inducible promoter responding to various types of stimuli to initiate transcription is used. In a preferred embodiment, the cell in which the toxin is expressed is a microorganism, such as a virus, bacteria, or a fungus. In a preferred embodiment, a virus, such as a baculovirus, contains a nucleotide sequence of the invention in its genome and expresses large amounts of the corresponding insecticidal toxin after infection of appropriate eukaryotic cells that are suitable for virus replication and expression of the nucleotide sequence. The insecticidal toxin thus produced is used as an insecticidal agent. Alternatively, baculoviruses engineered to include the nucleotide sequence are used to infect insects *in vivo* and kill them either by expression of the insecticidal toxin or by a combination of viral infection and expression of the insecticidal toxin.

Bacterial cells are also hosts for the expression of the nucleotide sequences of the invention. In a preferred embodiment, non-pathogenic symbiotic bacteria, which are able to live and replicate within plant tissues, so-called endophytes, or non- pathogenic symbiotic bacteria, which are capable of colonizing the phyllosphere or the rhizosphere, so-called epiphytes, are used. Such bacteria include bacteria of the genera *Agrobacterium, Alcaligenes, Azospirillum, Azotobacter, Bacillus, Clavibacter, Enterobacter, Erwinia, Flavobacter, Klebsiella, Pseudomonas, Rhizobium, Serratia, Streptomyces* and *Xanthomonas*. Symbiotic fungi, such as *Trichoderma* and *Gliocladium* are also possible hosts for expression of the inventive nucleotide sequences for the same purpose.

Techniques for these genetic manipulations are specific for the different available hosts and are known in the art. For example, the expression vectors pKK223-3 and pKK223-2 can be

used to express heterologous genes in *E. coli*, either in transcriptional or translational fusion, behind the tac or trc promoter. For the expression of operons encoding multiple ORFs, the simplest procedure is to insert the operon into a vector such as pKK223- 3 in transcriptional fusion, allowing the cognate ribosome binding site of the heterologous genes to be used. Techniques for overexpression in gram-positive species such as *Bacillus* are also known in the art and can be used in the context of this invention (Quax et al. In:Industrial Microorganisms:Basic and Applied Molecular Genetics, Eds. Baltz *et al.*, American Society for Microbiology, Washington (1993)). Alternate systems for overexpression rely for example, on yeast vectors and include the use of Pichia, Saccharomyces and Kluyveromyces (Sreekrishna, In:Industrial microorganisms:basic and applied molecular genetics, Baltz, Hegeman, and Skatrud eds., American Society for Microbiology, Washington (1993); Dequin & Barre, Biotechnology L2:173- 177 (1994); van den Berg *et al.*, Biotechnology 8:135-139 (1990)).

Plant transformation

In a particularly preferred embodiment, at least one of the insecticidal modified Cry3A toxins of the invention is expressed in a higher organism, e.g., a plant. In this case, transgenic plants expressing effective amounts of the modified Cry3A toxins protect themselves from insect pests. When the insect starts feeding on such a transgenic plant, it also ingests the expressed modified Cry3A toxins. This will deter the insect from further biting into the plant tissue or may even harm or kill the insect. A nucleotide sequence of the present invention is inserted into an expression cassette, which is then preferably stably integrated in the genome of said plant. In another preferred embodiment, the nucleotide sequence is included in a non-pathogenic self- replicating virus. Plants transformed in accordance with the present invention may be monocots or dicots and include, but are not limited to, maize, wheat, barley, rye, sweet potato, bean, pea, chicory, lettuce, cabbage, cauliflower, broccoli, turnip, radish, spinach, asparagus, onion, garlic, pepper, celery, squash, pumpkin, hemp, zucchini, apple, pear, quince, melon, plum, cherry, peach, nectarine, apricot, strawberry, grape, raspberry, blackberry, pineapple, avocado, papaya, mango, banana, soybean, tomato, sorghum, sugarcane, sugar

beet, sunflower, rapesced, clover, tobacco, carrot, cotton, alfalfa, rice, potato, eggplant, cucumber, Arabidopsis, and woody plants such as coniferous and deciduous trees.

Once a desired nucleotide sequence has been transformed into a particular plant species, it may be propagated in that species or moved into other varieties of the same species, particularly including commercial varieties, using traditional breeding techniques.

A nucleotide sequence of this invention is preferably expressed in transgenic plants, thus causing the biosynthesis of the corresponding modified Cry3A toxin in the transgenic plants. In this way, transgenic plants with enhanced resistance to insects are generated. For their expression in transgenic plants, the nucleotide sequences of the invention may require other modifications and optimization. Although in many cases genes from microbial organisms can be expressed in plants at high levels without modification, low expression in transgenic plants may result from microbial nucleotide sequences having codons that are not preferred in plants. It is known in the art that all organisms have specific preferences for codon usage, and the codons of the nucleotide sequences described in this invention can be changed to conform with plant preferences, while maintaining the amino acids encoded thereby. Furthermore, high expression in plants is best achieved from coding sequences that have at least about 35% GC content, preferably more than about 45%, more preferably more than about 50%, and most preferably more than about 60%. Microbial nucleotide sequences that have low GC contents may express poorly in plants due to the existence of ATTTA motifs that may destabilize messages, and AATAAA motifs that may cause inappropriate polyadenylation. Although preferred gene sequences may be adequately expressed in both monocotyledonous and dicotyledonous plant species, sequences can be modified to account for the specific codon preferences and GC content preferences of monocotyledons or dicotyledons as these preferences have been shown to differ (Murray *et al.* Nucl. Acids Res. 17:477-498 (1989)). In addition, the nucleotide sequences are screened for the existence of illegitimate splice sites that may cause message truncation. All changes required to be made within the nucleotide sequences such as those described above are made using well known techniques of site directed mutagenesis, PCR, and synthetic gene construction using the methods described in the published patent applications EP 0 385 962 (to Monsanto), EP 0 359 472 (to Lubrizol, and WO 93/07278 (to Ciba-Geigy).

In one embodiment of the invention a *cry*3A gene is made according to the procedure disclosed in U.S. Patent 5,625,136, herein incorporated by reference. In this procedure, maize preferred codons, i.e., the single codon that most frequently encodes that amino acid in maize, are used. The maize preferred codon for a particular amino acid might be derived, for example, from known gene sequences from maize. Maize codon usage for 28 genes from maize plants is found in Murray et al., Nucleic Acids Research 17:477-498 (1989), the disclosure of which is incorporated herein by reference. A synthetic sequence made with maize optimized codons is set forth in SEQ ID NO: 3.

In this manner, the nucleotide sequences can be optimized for expression in any plant. It is recognized that all or any part of the gene sequence may be optimized or synthetic. That is, synthetic or partially optimized sequences may also be used.

For efficient initiation of translation, sequences adjacent to the initiating methionine may require modification. For example, they can be modified by the inclusion of sequences known to be effective in plants. Joshi has suggested an appropriate consensus for plants (NAR 15:6643-6653 (1987)) and Clonetech suggests a further consensus translation initiator (1993/1994 catalog, page 210). These consensuses are suitable for use with the nucleotide sequences of this invention. The sequences are incorporated into constructions comprising the nucleotide sequences, up to and including the ATG (whilst leaving the second amino acid unmodified), or alternatively up to and including the GTC subsequent to the ATG (with the possibility of modifying the second amino acid of the transgene).

Expression of the nucleotide sequences in transgenic plants is driven by promoters that function in plants. The choice of promoter will vary depending on the temporal and spatial requirements for expression, and also depending on the target species. Thus, expression of the nucleotide sequences of this invention in leaves, in stalks or stems, in ears, in inflorescences (e.g. spikes, panicles, cobs, etc.), in roots, and/or seedlings is preferred. In many cases, however, protection against more than one type of insect pest is sought, and thus expression in multiple tissues is desirable. Although many promoters from dicotyledons have been shown to be operational in monocotyledons and vice versa, ideally dicotyledonous promoters are selected for expression in dicotyledons, and monocotyledonous promoters for expression in monocotyledons. However, there is no restriction to the provenance of selected promoters; it

is sufficient that they are operational in driving the expression of the nucleotide sequences in the desired cell.

Preferred promoters that are expressed constitutively include promoters from genes encoding actin or ubiquitin and the CaMV 35S and 19S promoters. The nucleotide sequences of this invention can also be expressed under the regulation of promoters that are chemically regulated. This enables the insecticidal modified Cry3A toxins to be synthesized only when the crop plants are treated with the inducing chemicals. Preferred technology for chemical induction of gene expression is detailed in the published application EP 0 332 104 (to Ciba-Geigy) and U.S. Patent 5,614,395. A preferred promoter for chemical induction is the tobacco PR-1a promoter.

A preferred category of promoters is that which is wound inducible. Numerous promoters have been described which are expressed at wound sites and also at the sites of phytopathogen infection. Ideally, such a promoter should only be active locally at the sites of infection, and in this way the insecticidal modified Cry3A toxins only accumulate in cells that need to synthesize the insecticidal modified Cry3A toxins to kill the invading insect pest. Preferred promoters of this kind include those described by Stanford *et al.* Mol. Gen. Genet. 215:200-208 (1989), Xu *et al.* Plant Molec. Biol. 22:573-588 (1993), Logemann *et al.* Plant Cell 1:151-158 (1989), Rohrmeier & Lehle, Plant Molec. Biol. 22:783-792 (1993), Firek *et al.* Plant Molec. Biol. 22:129-142 (1993), and Warner *et al.* Plant J. 3:191-201 (1993).

Tissue-specific or tissue-preferential promoters useful for the expression of the modified Cry3A toxin genes in plants, particularly maize, are those which direct expression in root, pith, leaf or pollen, particularly root. Such promoters, e.g. those isolated from PEPC or trpA, are disclosed in U.S. Pat. No. 5,625,136, or MTL, disclosed in U.S. Pat. No. 5,466,785. Both U. S. patents are herein incorporated by reference in their entirety.

Further preferred embodiments are transgenic plants expressing the nucleotide sequences in a wound-inducible or pathogen infection-inducible manner.

In addition to promoters, a variety of transcriptional terminators are also available for use in chimeric gene construction using the modified Cry3A toxin genes of the present invention. Transcriptional terminators are responsible for the termination of transcription beyond the transgene and its correct polyadenylation. Appropriate transcriptional terminators and those that are known to function in plants include the CaMV 35S terminator, the tml terminator, the

nopaline synthase terminator, the pea rbcS E9 terminator and others known in the art. These can be used in both monocotyledons and dicotyledons. Any available terminator known to function in plants can be used in the context of this invention.

Numerous other sequences can be incorporated into expression cassettes described in this invention. These include sequences that have been shown to enhance expression such as intron sequences (e.g. from Adh1 and bronze1) and viral leader sequences (e.g. from TMV, MCMV and AMV).

It may be preferable to target expression of the nucleotide sequences of the present invention to different cellular localizations in the plant. In some cases, localization in the cytosol may be desirable, whereas in other cases, localization in some subcellular organelle may be preferred. Subcellular localization of transgene-encoded enzymes is undertaken using techniques well known in the art. Typically, the DNA encoding the target peptide from a known organelle-targeted gene product is manipulated and fused upstream of the nucleotide sequence. Many such target sequences are known for the chloroplast and their functioning in heterologous constructions has been shown. The expression of the nucleotide sequences of the present invention is also targeted to the endoplasmic reticulum or to the vacuoles of the host cells. Techniques to achieve this are well known in the art.

Vectors suitable for plant transformation are described elsewhere in this specification. For Agrobacterium-mediated transformation, binary vectors or vectors carrying at least one T-DNA border sequence are suitable, whereas for direct gene transfer any vector is suitable and linear DNA containing only the construction of interest may be preferred. In the case of direct gene transfer, transformation with a single DNA species or co-transformation can be used (Schocher *et al.* Biotechnology 4:1093- 1096 (1986)). For both direct gene transfer and Agrobacterium-mediated transfer, transformation is usually (but not necessarily) undertaken with a selectable marker that may provide resistance to an antibiotic (kanamycin, hygromycin or methotrexate) or a herbicide (basta). Plant transformation vectors comprising the modified Cry3A toxin genes of the present invention may also comprise genes (e.g. phosphomannose isomerase; PMI) which provide for positive selection of the transgenic plants as disclosed in U.S. Patents 5,767,378 and 5,994,629, herein incorporated by reference. The choice of selectable marker is not, however, critical to the invention.

In another embodiment, a nucleotide sequence of the present invention is directly transformed into the plastid genome. A major advantage of plastid transformation is that plastids are generally capable of expressing bacterial genes without substantial codon optimization, and plastids are capable of expressing multiple open reading frames under control of a single promoter. Plastid transformation technology is extensively described in U.S. Patent Nos. 5,451,513, 5,545,817, and 5,545,818, in PCT application no. WO 95/16783, and in McBride *et al.* (1994) Proc. Nati. Acad. Sci. USA 91, 7301-7305. The basic technique for chloroplast transformation involves introducing regions of cloned plastid DNA flanking a selectable marker together with the gene of interest 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. Nati. Acad. Sci. USA 87, 8526-8530; Staub, J. M., and Maliga, P. (1992) Plant Cell 4, 39-45). This resulted in stable homoplasmic transformants at a frequency of approximately one per 100 bombardments of target leaves. 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-cletoxifying enzyme aminoglycoside- 3'- adenyltransferase (Svab, Z., and Maliga, P. (1993) Proc. Natl. Acad. Sci. USA 90, 913-917). Previously, this marker had been used successfully for high-frequency transformation of the plastid genome of the green alga Chlamydomonas reinhardtii (Goldschmidt- Clermont, M. (1991) Nucl. Acids Res. 19:4083-4089). Other selectable markers useful for plastid transformation are known in the art and encompassed within the scope of the invention. Typically, approximately 15-20 cell division cycles following transformation are required to reach a homoplastidic state. Plastid expression, in which genes are inserted by homologous recombination into all of 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, a nucleotide sequence of the present invention is inserted into a plastid-targeting vector and transformed into the plastid genome of a desired plant host. Plants homoplastic for plastid genomes containing a nucleotide sequence of the present invention are obtained, and are preferentially capable of high expression of the nucleotide sequence.

Combinations of Insect Control Principles

The modified Cry3A toxins of the invention can be used in combination with Bt δ-endotoxins or other pesticidal principles to increase pest target range. Furthermore, the use of the modified Cry3A toxins of the invention in combination with Bt δ-endotoxins or other pesticidal principles of a distinct nature has particular utility for the prevention and/or management of insect resistance.

Other insecticidal principles include, for example, lectins, α-amylase, peroxidase and cholesterol oxidase. Vegetative Insecticidal Protein genes, such as *vip*1A(a) and *vip*2A(a) as disclosed in U.S. Pat. No. 5,889,174 and herein incorporated by reference, are also useful in the present invention.

This co-expression of more than one insecticidal principle in the same transgenic plant can be achieved by genetically engineering a plant to contain and express all the genes necessary. Alternatively, a plant, Parent 1, can be genetically engineered for the expression of genes of the present invention. A second plant, Parent 2, can be genetically engineered for the expression of a supplemental insect control principle. By crossing Parent 1 with Parent 2, progeny plants are obtained which express all the genes introduced into Parents 1 and 2.

Transgenic seed of the present invention can also be treated with an insecticidal seed coating as described in U. S. Patent Nos. 5,849,320 and 5,876,739, herein incorporated by reference. Where both the insecticidal seed coating and the transgenic seed of the invention are active against the same target insect, the combination is useful (i) in a method for enhancing activity of a modified Cry3A toxin of the invention against the target insect and (ii) in a method for preventing development of resistance to a modified Cry3A toxin of the invention by providing a second mechanism of action against the target insect. Thus, the invention provides a method

of enhancing activity against or preventing development of resistance in a target insect, for example corn rootworm, comprising applying an insecticidal seed coating to a transgenic seed comprising one or more modified Cry3A toxins of the invention.

Even where the insecticidal seed coating is active against a different insect, the insecticidal seed coating is useful to expand the range of insect control, for example by adding an insecticidal seed coating that has activity against lepidopteran insects to the transgenic seed of the invention, which has activity against coleopteran insects, the coated transgenic seed produced controls both lepidopteran and coleopteran insect pests.

EXAMPLES

The invention will be further described by reference to the following detailed examples. These examples are provided for the purposes of illustration only, and are not intended to be limiting unless otherwise specified. Standard recombinant DNA and molecular cloning techniques used here are well known in the art and are described by J. Sambrook, *et al.*, Molecular Cloning: *A Laboratory Manual, 3d Ed.*, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press (2001); by T.J. Silhavy, M.L. Berman, and L.W. Enquist, *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) and by Ausubel, F.M. *et al., Current Protocols in Molecular Biology*, New York, John Wiley and Sons Inc., (1988), Reiter, *et al., Methods in Arabidopsis Research*, World Scientific Press (1992), and Schultz *et al.*, *Plant Molecular Biology Manual*, Kluwer Academic Publishers (1998).

Example 1: Maize Optimized *cry*3A Gene Construction

The maize optimized *cry*3A gene is made according to the procedure disclosed in U.S. Patent 5,625,136. In this procedure, maize preferred codons, i.e., the single codon that most frequently encodes that amino acid in maize, are used. The maize preferred codon for a particular amino acid is derived from known gene sequences from maize. Maize codon usage for 28 genes from maize plants is found in Murray *et al.*, Nucleic Acids Research 17:477-498 (1989). The synthetic sequence made with maize optimized codons is set forth in SEQ ID NO: 3.

Example 2: **Identification of Cathepsin-G Enzymatic Activity in Western Corn Rootworm Guts**

Cathepsin G-like (serine protease) and cathepsin B-like (cysteine protease) enzymatic activities in western corn rootworm guts are measured using colorimetric substrates. Each 1 ml reaction contains five homogenized midguts of the 3rd instar of western corn rootworm and 1 mg of substrate dissolved in reaction buffer (10 mM Tris, 5 mM NaCl, 0.01 M DTT, pH 7.5). The cathepsin G substrate tested is Ala-Ala-Pro-Phe (SEQ ID NO: 35)-*p*NA and cathepsin B substrate, Arg-Arg-*p*NA. The reactions are incubated at 28°C for 1hr. The intensity of yellow color formation, indicative of the efficiency of a protease to recognize the appropriate substrate, is compared in treatments vs. controls. The reactions are scored as negative (-) if no color or slight background color is detected. Reactions which are 25%, 50%, 75% or 100% above background are scored as +, ++, +++, or ++++, respectively. Results of the enzymatic assays are shown in the following table.

Table 1

Reaction	Product Color intensity
WCR gut only	-
Cathepsin B substrate only	-
Cathepsin G substrate only	-
WCR gut + Cathepsin B substrate	+
WCR gut + Cathepsin G substrate	+++

This is the first time that the serine protease cathepsin G activity has been identified in western corn rootworm guts. Western corn rootworm guts clearly have stronger cathepsin G, the serine protease, activity compared to cathepsin B, the cysteine protease, activity. The AAPF sequence (SEQ ID NO: 35) is selected as the cathepsin G protease recognition site for creating modified Cry3A toxins of the present invention.

Example 3: **Construction of Modified *cry*3A Genes**

Modified *cry*3A genes comprising a nucleotide sequence that encodes the cathepsin G recognition site in domain I, domain III, or domain I and domain III are made using overlap

PCR. The maize optimized *cry*3A gene (SEQ ID NO: 2), comprised in plasmid pCIB6850 (SEQ ID NO: 5), is used as the starting template. Eight modified *cry*3A gene constructs, which encode modified Cry3A toxins, are made; *cry*3A054, *cry*3A055, and *cry*3A085, which comprise the cathepsin G recognition site coding sequence in domain I; *cry*3A058, *cry*3A082, which comprise the cathepsin G recognition site coding sequence in domain III; *cry*3A056, *cry*3A057, *cry*3A083, which comprise the cathepsin G recognition site coding sequence in domain I and domain III. The eight modified *cry*3A genes and the modified Cry3A toxins they encode are described as follows:

*cry*3A054 comprised in pCMS054

*cry*3A054 (SEQ ID NO: 6) comprises a nucleotide sequence encoding a modified Cry3A toxin. Three overlap PCR primer pairs are used to insert the nucleotide sequence encoding the cathepsin G recognition site into the unmodified maize optimized *cry*3A:

1. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3' (SEQ ID NO: 22)
AAPFtail3 – 5'-GAACGGTGCAGCGGGGTTCTTCTGCCAGC-3' (SEQ ID NO: 23)

2. Tail5mod – 5'-GCTGCACCGTTCCCCCACAGCCAGGGCCG-3' (SEQ ID NO: 24)
XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3' (SEQ ID NO: 25)

3. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3' (SEQ ID NO: 22)
XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3' (SEQ ID NO: 25)

Primer pair 1 and primer pair 2 generate two unique PCR products. These products are then combined in equal parts and primer pair 3 is used to join the products to generate one PCR fragment that is cloned back into the original pCIB6850 template. The modified *cry*3A054 gene is then transferred to pBluescript (Stratagene). The resulting plasmid is designated pCMS054 and comprises the *cry*3A054 gene (SEQ ID NO: 6).

The modified Cry3A054 toxin (SEQ ID NO: 7), encoded by the modified *cry*3A gene comprised in pCMS054, has a cathepsin G recognition site, comprising the amino acid sequence AAPF (SEQ ID NO: 35), inserted in domain I between amino acids 107 and 113 of

the unmodified Cry3A toxin (SEQ ID NO: 4). The cathepsin G recognition site replaces the naturally occurring trypsin recognition site and is adjacent to a naturally occurring chymotrypsin recognition site.

*cry*3A055 comprised in pCMS055

*cry*3A055 (SEQ ID NO: 8) comprises a nucleotide sequence encoding a modified Cry3A toxin. Three overlap PCR primer pairs are used to insert the nucleotide sequence encoding the cathepsin G recognition site into the unmodified maize optimized *cry*3A:

1. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3' (SEQ ID NO: 22)
AAPFtail3 – 5'-GAACGGTGCAGCGGGGTTCTTCTGCCAGC-3' (SEQ ID NO: 23)

2. AAPFtail4 – 5'-GCTGCACCGTTCCGCAACCCCCACAGCCA-3' (SEQ ID NO: 26)
XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3' (SEQ ID NO: 25)

3. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3' (SEQ ID NO: 22)
XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3' (SEQ ID NO: 25)

Primer pair 1 and primer pair 2 generate two unique PCR products. These products are then combined in equal parts and primer pair 3 is used to join the products to generate one PCR fragment that is cloned back into the original pCIB6850 template. The modified *cry*3A055 gene is then transferred to pBluescript (Stratagene). The resulting plasmid is designated pCMS055 and comprises the *cry*3A055 gene (SEQ ID NO: 8).
The modified Cry3A055 toxin (SEQ ID NO: 9), encoded by the modified *cry*3A gene comprised in pCMS055, has a cathepsin G recognition site, comprising the amino acid sequence AAPF (SEQ ID NO: 35), inserted in domain I between amino acids 107 and 111 of the unmodified Cry3A toxin (SEQ ID NO: 4). The cathepsin G recognition site is adjacent to a natural trypsin and chymotrypsin recognition site.

*cry*3A058 comprised in pCMS058

*cry*3A058 (SEQ ID NO: 14) comprises a nucleotide sequence encoding a modified Cry3A toxin. Three overlap PCR primer pairs are used to insert the nucleotide sequence encoding the cathepsin G recognition site into the unmodified maize optimized *cry*3A:

1. SalExt – 5'-GAGCGTCGACTTCTTCAAC-3' (SEQ ID NO: 27)
AAPF-Y2 – 5'-GAACGGTGCAGCGTATTGGTTGAAGGGGGC-3' (SEQ ID NO: 28)

2. AAPF-Y1 – 5'-GCTGCACCGTTCTACTTCGACAAGACCATC-3' (SEQ ID NO: 29)
SacExt – 5'-GAGCTCAGATCTAGTTCACGG-3' (SEQ ID NO: 30)

3. SalExt – 5'-GAGCGTCGACTTCTTCAAC-3' (SEQ ID NO: 27)
SacExt – 5'-GAGCTCAGATCTAGTTCACGG-3' (SEQ ID NO: 30)

Primer pair 1 and primer pair 2 generate two unique PCR products. These products are then combined in equal parts and primer pair 3 is used to join the products to generate one PCR fragment that is cloned back into the original pCIB6850 template. The modified *cry*3A058 gene is then transferred to pBluescript (Stratagene). The resulting plasmid is designated pCMS058 and comprises the *cry*3A058 gene (SEQ ID NO: 14).

The modified Cry3A058 toxin (SEQ ID NO: 15), encoded by the modified *cry*3A gene, has a cathepsin G recognition site, comprising the amino acid sequence AAPF (SEQ ID NO: 35), inserted in domain III between amino acids 540and 541 of the unmodified Cry3A toxin (SEQ ID NO: 4). The cathepsin G recognition site is within a naturally occurring chymotrypsin recognition site.

pCMS082 comprising *cry*3A082

*cry*3A082 (SEQ ID NO: 12) comprises a nucleotide sequence encoding a modified Cry3A toxin. A QuikChange Site Directed Mutagenesis PCR primer pair is used to insert the nucleotide sequence encoding the cathepsin G recognition site into the unmodified maize optimized *cry*3A:

BBmod1 – 5'-CGGGGCCCCCGCTGCACCGTTCTACTTCGACA-3' (SEQ ID NO: 31)
BBmod2 – 5'-TGTCGAAGTAGAACGGTGCAGCGGGGGCCCCG-3' (SEQ ID NO: 32)

The primer pair generates a unique PCR product. This product is cloned back into the original pCIB6850 template. The modified *cry*3A082 gene is then transferred to pBluescript (Stratagene). The resulting plasmid is designated pCMS082 and comprises the *cry*3A082 gene (SEQ ID NO: 12).

The modified Cry3A082 toxin (SEQ ID NO: 13), encoded by the modified *cry*3A gene, has a cathepsin G recognition site, comprising the amino acid sequence AAPF (SEQ ID NO: 35), inserted in domain III between amino acids 539 and 542 of the unmodified Cry3A toxin (SEQ ID NO: 4). The cathepsin G recognition site replaces a naturally occurring chymotrypsin recognition site.

<u>*cry*3A056 comprised in pCMS056</u>

*cry*3A056 (SEQ ID NO: 18) comprises a nucleotide sequence encoding a modified Cry3A toxin. Six overlap PCR primer pairs are used to insert two cathepsin G recognition sites into the unmodified *cry*3A:

1. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3' (SEQ ID NO: 22)
AAPFtail3 – 5'-GAACGGTGCAGCGGGGTTCTTCTGCCAGC-3' (SEQ ID NO: 23)

2. AAPFtail4 – 5'-GCTGCACCGTTCCGCAACCCCCACAGCCA-3' (SEQ ID NO: 26)
XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3' (SEQ ID NO: 25)

3. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3' (SEQ ID NO: 22)
XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3' (SEQ ID NO: 25)

4. SalExt – 5'-GAGCGTCGACTTCTTCAAC-3' (SEQ ID NO: 27)
AAPF-Y2 – 5'-GAACGGTGCAGCGTATTGGTTGAAGGGGGC-3' (SEQ ID NO: 28)

5. AAPF-Y1 – 5'-GCTGCACCGTTCTACTTCGACAAGACCATC-3' (SEQ ID NO: 29)
SacExt – 5'-GAGCTCAGATCTAGTTCACGG-3' (SEQ ID NO: 30)

6. SalExt – 5'-GAGCGTCGACTTCTTCAAC-3' (SEQ ID NO: 27)
SacExt – 5'-GAGCTCAGATCTAGTTCACGG-3' (SEQ ID NO: 30)

Primer pair 1 and primer pair 2 generate two unique PCR products. These products are combined in equal parts and primer pair 3 is used to join the products to generate one PCR fragment that is cloned back into the original pCIB6850 plasmid. The modified *cry*3A055 gene is then transferred to pBluescript (Stratagene). The resulting plasmid is designated pCMS055. Primer pair 4 and primer pair 5 generate another unique set of fragments that are joined by another PCR with primer pair 6. This fragment is cloned into domain III of the modified *cry*3A055 gene comprised in pCMS055. The resulting plasmid is designated pCMS056 and comprises the *cry*3A056 gene (SEQ ID NO: 18).

The modified Cry3A056 toxin (SEQ ID NO: 19), encoded by the modified *cry*3A gene, has a cathepsin G recognition site, comprising the amino acid sequence AAPF (SEQ ID NO: 35), inserted in domain I between amino acids 107 and 111 and in domain III between amino acids 540 and 541 of the unmodified Cry3A toxin (SEQ ID NO: 4). The cathepsin G recognition site is adjacent to a naturally occurring trypsin and chymotrypsin recognition site in domain I and is within a naturally occurring chymotrypsin recognition site in domain III.

*cry*3A057 comprised in pCMS057

*cry*3A057 (SEQ ID NO: 16) comprises a nucleotide sequence encoding a modified Cry3A toxin. Six overlap PCR primer pairs are used to insert two cathepsin G recognition sites into the unmodified *cry*3A:

1. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3' (SEQ ID NO: 22)
AAPFtail3 – 5'-GAACGGTGCAGCGGGGTTCTTCTGCCAGC-3' (SEQ ID NO: 23)

2. Tail5mod – 5'-GCTGCACCGTTCCCCCACAGCCAGGGCCG-3' (SEQ ID NO: 24)
XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3' (SEQ ID NO: 25)

3. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3' (SEQ ID NO: 22)
XbalExt2 – 5'-TCTAGACCCACGTTGTACCAC-3' (SEQ ID NO: 25)

4. SalExt – 5'-GAGCGTCGACTTCTTCAAC-3' (SEQ ID NO: 27)
AAPF-Y2 – 5'-GAACGGTGCAGCGTATTGGTTGAAGGGGGC-3' (SEQ ID NO: 28)

5. AAPF-Y1 – 5'-GCTGCACCGTTCTACTTCGACAAGACCATC-3' (SEQ ID NO: 29)
SacExt – 5'-GAGCTCAGATCTAGTTCACGG-3' (SEQ ID NO: 30)

6. SalExt – 5'-GAGCGTCGACTTCTTCAAC-3' (SEQ ID NO: 27)
SacExt – 5'-GAGCTCAGATCTAGTTCACGG-3' (SEQ ID NO: 30)

Primer pair 1 and primer pair 2 generate two unique PCR products. These products are combined in equal parts and primer pair 3 is used to join the products to generate one PCR fragment that is cloned back into the original pCIB6850 plasmid. The modified *cry*3A054 gene is then transferred to pBluescript (Stratagene). The resulting plasmid is designated pCMS054. Primer pair 4 and primer pair 5 generate another unique set of fragments that are joined by another PCR with primer pair 6. This fragment is cloned into domain III of the modified *cry*3A054 gene comprised in pCMS054. The resulting plasmid is designated pCMS057 and comprises the *cry*3A057 gene (SEQ ID NO: 16).

The modified Cry3A057 toxin (SEQ ID NO: 17), encoded by the modified *cry*3A gene, has a cathepsin G recognition site, comprising the amino acid sequence AAPF (SEQ ID NO: 35), inserted in domain I between amino acids 107 and 113 and in domain III between amino acids 540 and 541 of the unmodified Cry3A toxin (SEQ ID NO: 4). The cathepsin G recognition site replaces a naturally occurring trypsin recognition site and is adjacent to a naturally occurring chymotrypsin recognition site in domain I and is within a naturally occurring chymotrypsin recognition site in domain III.

*cry*3A083 comprised in pCMS083

*cry*3A083 (SEQ ID NO: 20) comprises a nucleotide sequence encoding a modified Cry3A toxin. Three overlap PCR primer pairs and one QuikChange Site Directed Mutagenesis PCR primer pair are used to insert two cathepsin G recognition sites into the unmodified *cry*3A:

1. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3'	(SEQ ID NO: 22)
AAPFtail3 – 5'-GAACGGTGCAGCGGGGTTCTTCTGCCAGC-3'	(SEQ ID NO: 23)
2. AAPFtail4 – 5'-GCTGCACCGTTCCGCAACCCCCACAGCCA-3'	(SEQ ID NO: 26)
XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3'	(SEQ ID NO: 25)
3. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3'	(SEQ ID NO: 22)
XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3'	(SEQ ID NO: 25)
BBmod1 – 5'-CGGGGCCCCCGCTGCACCGTTCTACTTCGACA-3	(SEQ ID NO: 31)
BBmod2 – 5'-TGTCGAAGTAGAACGGTGCAGCGGGGGCCCCG-3'	(SEQ ID NO: 32)

Primer pair 1 and primer pair 2 generate two unique PCR products. These products are combined in equal parts and primer pair 3 is used to join the products to generate one PCR fragment that is cloned back into the original pCIB6850 plasmid. The modified *cry*3A055 gene is then transferred to pBluescript (Stratagene). The resulting plasmid is designated pCMS055. Primer pair 4 generates another unique fragment that is cloned into domain III of the modified *cry*3A comprised in pCMS055. The resulting plasmid is designated pCMS083 and comprises the *cry*3A083 gene (SEQ ID NO: 20).

The modified Cry3A083 toxin (SEQ ID NO: 21), encoded by the modified *cry*3A gene, has a cathepsin G recognition site, comprising the amino acid sequence AAPF (SEQ ID NO: 35), inserted in domain I between amino acids 107 and 111 and between amino acids 539 and 542 of the unmodified Cry3A toxin (SEQ ID NO: 4). The cathepsin G recognition site is adjacent to a naturally occurring trypsin and chymotrypsin recognition site in domain I and replaces a naturally occurring chymotrypsin recognition site in domain III.

*cry*3A085 comprised in pCMS085

The *cry*3A085 gene (SEQ ID NO: 10) comprises a cathepsin G coding sequence at the same position as in the *cry*3A055 gene described above. The *cry*3A085 gene has an additional 24 nucleotides inserted at the 5' end which encode amino acids 41-47 of the deduced amino acid sequence set forth in SEQ ID NO: 2 as well as an additional methionine. The additional nucleotides are inserted at the 5' end of the *cry*3A055 gene using the following PCR primer pair:

mo3Aext- 5'-GGATCCACCATGAACTACAAGGAGTTCCTCCGC-ATGACCGCCGACAAC-3' (SEQ ID NO: 33)

CMS16 – 5'-CCTCCACCTGCTCCATGAAG-3' (SEQ ID NO: 34)

The modified Cry3A085 toxin (SEQ ID NO: 11), encoded by the modified *cry*3A gene, has a cathepsin G recognition site, comprising the amino acid sequence AAPF (SEQ ID NO: 35), inserted in domain I between amino acids corresponding to 107 and 111 of the unmodified Cry3A toxin (SEQ ID NO: 4) and has an additional eight amino acid residues at the N-terminus of which the second residue corresponds to amino acid number 41 of the amino acid sequence set forth in SEQ ID NO: 2.

Example 4: **Insecticidal Activity of Modified Cry3A Toxins**

Modified Cry3A toxins are tested for insecticidal activity against western corn rootworm, northern corn rootworm and southern corn rootworm in insect bioassays. Bioassays are performed using a diet incorporation method. *E. coli* clones that express one of the modified Cry3A toxins of the invention are grown overnight. 500 µl of an overnight culture is sonicated and then mixed with 500 µl of molten artificial diet (Marrone et al. (1985) J. of Economic Entomology 78:290-293). Once the diet solidifies, it is dispensed in a petri-dish and 20 neonate corn rootworm are placed on the diet. The petri-dishes are held at 30°C. Mortality is recorded after 6 days. All of the modified Cry3A toxins cause 50%-100% mortality to western and northern corn rootworm whereas the unmodified Cry3A toxin causes 0%-30% mortality. None of the modified Cry3A toxins have activity against southern corn rootworm.

Example 5: **Creation of Transgenic Maize Plants Comprising Modified *cry*3A Coding Sequences**

Three modified cry3A genes, *cry*3A055, representative of a domain I modification, *cry*3A058, representative of a domain III modification, and *cry*3A056, representative of a domain I and domain III modification, are chosen for transformation into maize plants. An expression cassette comprising a modified *cry*3A coding sequence is transferred to a suitable vector for Agrobacterium-mediated maize transformation. For this example, an expression cassette comprises, in addition to the modified *cry*3A gene, the MTL promoter (U.S. Pat. No. 5,466,785) and the nos terminater which is known in the art.

Transformation of immature maize embryos is performed essentially as described in Negrotto *et al.*, 2000, Plant Cell Reports 19: 798-803. For this example, all media constituents are as described in Negrotto *et al.*, *supra*. However, various media constituents known in the art may be substituted.

The genes used for transformation are cloned into a vector suitable for maize transformation. Vectors used in this example contain the phosphomannose isomerase (PMI) gene for selection of transgenic lines (Negrotto *et al.* (2000) Plant Cell Reports 19: 798-803).

Agrobacterium strain LBA4404 (pSB1) containing the plant transformation plasmid is grown on YEP (yeast extract (5 g/L), peptone (10g/L), NaCl (5g/L), 15g/l agar, pH 6.8) solid medium for 2 – 4 days at 28°C. Approximately 0.8X 10^9 *Agrobacterium* are suspended in LS-inf media supplemented with 100 µM As (Negrotto *et al.*,(2000) Plant Cell Rep 19: 798-803). Bacteria are pre-induced in this medium for 30-60 minutes.

Immature embryos from A188 or other suitable genotype are excised from 8 – 12 day old ears into liquid LS-inf + 100 µM As. Embryos are rinsed once with fresh infection medium. *Agrobacterium* solution is then added and embryos are vortexed for 30 seconds and allowed to settle with the bacteria for 5 minutes. The embryos are then transferred scutellum side up to LSAs medium and cultured in the dark for two to three days. Subsequently, between 20 and 25 embryos per petri plate are transferred to LSDc medium supplemented with cefotaxime (250 mg/l) and silver nitrate (1.6 mg/l) and cultured in the dark for 28°C for 10 days.

Immature embryos, producing embryogenic callus are transferred to LSD1M0.5S medium. The cultures are selected on this medium for 6 weeks with a subculture step at 3 weeks.

Surviving calli are transferred to Reg1 medium supplemented with mannose. Following culturing in the light (16 hour light/ 8 hour dark regiment), green tissues are then transferred to Reg2 medium without growth regulators and incubated for 1-2 weeks. Plantlets are transferred to Magenta GA-7 boxes (Magenta Corp, Chicago Ill.) containing Reg3 medium and grown in the light. After 2-3 weeks, plants are tested for the presence of the PMI genes and the modified cry3A genes by PCR. Positive plants from the PCR assay are transferred to the greenhouse and tested for resistance to corn rootworm.

Example 6: **Analysis of Transgenic Maize Plants**

Corn Rootworm Efficacy

Root Excision Bioassay

Plants are sampled as they are being transplanted from Magenta GA-7 boxes into soil. This allows the roots to be sampled from a reasonably sterile environment relative to soil conditions. Sampling consists of cutting a small piece of root (ca. 2-4 cm long) and placing it onto enriched phytagar (phytagar, 12 g., sucrose, 9 g., MS salts, 3 ml., MS vitamins, 3 ml., Nystatin(25mg/ml), 3 ml., Cefotaxime (50mg/ml), 7 ml., Aureomycin (50 mg/ml), 7 ml., Streptomycin (50mg/ml), 7 ml., dH_2O, 600 ml) in a small petri-dish. Negative controls are either transgenic plants that are PCR negative for the modified *cry*3A gene from the same experiment, or from non-transgenic plants (of a similar size to test plants) that are being grown in the phytotron. If sampling control roots from soil, the root samples are washed with water to remove soil residue, dipped in Nystatin solution (5mg/ml), removed from the dip, blotted dry with paper toweling, and placed into a phytagar dish.

Root samples are inoculated with western corn rootworms by placing 10 first instar larvae onto the inside surface of the lid of each phytagar dish and the lids then tightly resealed. Larvae are handled using a fine tip paintbrush. After all dishes are inoculated, the tray of dishes is placed in the dark at room temperature until data collection.

At 3-4 days post inoculation, data is collected. The percent mortality of the larvae is calculated along with a visual damage rating of the root. Feeding damage is rated as high, moderate, low, or absent and given a numerical value of 3, 2, 1 or 0, respectively. Root

samples causing at least 40% mortality and having a damage rating of 2 or less are considered positive.

Results in the following table show that plants expressing a modified Cry3A toxin cause from 40-100% mortality to western corn rootworm whereas control plants cause 0-30% mortality. Also, plants expressing a modified Cry3A toxin sustain significantly less feeding damage than control plants.

Table 2

T0 Event	Modified Cry3A Toxin Expressed	Percent Mortality Per Plant A	B	C	D	E	Mean Damage Rating Per Event
240A7	Cry3A055	80	40	80	60		0.8
240B2	Cry3A055	60	60	60	80		1.25
240B9	Cry3A055	40	60	60	100		1
240B10	Cry3A055	80	40	60	60		1
240A15	Cry3A055	80	60	50	70	70	0.6
240A5	Cry3A055	60	80	60			0.33
240A9	Cry3A055	50	60	60	70	70	1.6
244A4	Cry3A058	50					1
244A7	Cry3A058	40	40	60			1.3
244A5	Cry3A058	50					1
244B7	Cry3A058	90					1
244B6	Cry3A058	50	40	60			1
243A3	Cry3A056	50	90	80	60		1.25
243A4	Cry3A056	50	80	60			1.7
243B1	Cry3A056	80	90				0.5
243B4	Cry3A056	70	60	50	80		1.5
245B2	Cry3A056	90	50	70	60		1
WT1	-	0	10	20	10	0	2.6
WT2	-	0	30	0	0	20	2.8

Whole Plant Bioassay

Some positive plants identified using the root excision bioassay described above are evaluated for western corn rootworm resistance using a whole plant bioassay. Plants are infested generally within 3 days after the root excision assay is completed.

Western corn rootworm eggs are preincubated so that hatch occurs 2-3 days after plant inoculation. Eggs are suspended in 0.2% agar and applied to the soil around test plants at approximately 200 eggs/plant.

Two weeks after the eggs hatch, plants are evaluated for damage caused by western corn rootworm larvae. Plant height attained, lodging, and root mass are criteria used to determine if plants are resistant to western corn rootworm feeding damage. At the time of evaluation, control plants typically are smaller than modified Cry3A plants. Also, non-transgenic control plants and plants expressing the unmodified Cry3A toxin encoded by the maize optimized *cry*3A gene have lodged during this time due to severe pruning of most of the roots resulting in no root mass accumulation. At the time of evaluation, plants expressing a modified Cry3A toxin of the invention are taller than control plants, have not lodged, and have a large intact root mass due to the insecticidal activity of the modified Cry3A toxin.

ELISA Assay

ELISA analysis according to the method disclosed in U.S. Patent No. 5,625,136 is used for the quantitative determination of the level of modified and unmodified Cry3A protein in transgenic plants.

Table 3: Whole Plant Bioassay Results and Protein Levels

Transgenic Maize Plant	Type of Cry3A Toxin Expressed	Cry3A Protein Level in Roots (ng/mg)	Plant Lodged	Intact Root Mass
240A2E	modified Cry3A055	224	-	+
240A9C	modified Cry3A055	71	-	+
240B9D	modified Cry3A055	204	-	+
240B9E	modified Cry3A055	186	-	+
240B10D	modified Cry3A055	104	-	+
240B10E	modified Cry3A055	70	-	+
240A15E	modified Cry3A055	122	-	+
240B4D	modified Cry3A055	97	-	+
243B5A	modified Cry3A056	41	-	+
244A7A	modified Cry3A058	191	-	+
710-2-51	maize optimized	39	+	-
710-2-54	maize optimized	857	+	-
710-2-61	maize optimized	241	+	-
710-2-67	maize optimized	1169	+	-
710-2-68	maize optimized	531	+	-
710-2-79	maize optimized	497	+	-
710-2-79	maize optimized	268	+	-
WT1 Control	-	0	+	-
WT2 Control	-	0	+	-

What is claimed is:

1. An isolated nucleic acid molecule comprising a nucleotide sequence that encodes a modified Cry3A toxin comprising a non-naturally occurring protease recognition site, wherein said protease recognition site modifies a Cry3A toxin and is located at a position selected from the group consisting of:
   a) a position between amino acids corresponding to amino acid numbers 107 and 115 of SEQ ID NO:4;
   b) a position between amino acids corresponding to amino acid numbers 536 and 542 of SEQ ID NO:4; and
   c) a position between amino acids corresponding to amino acid numbers 107 and 115 of SEQ ID NO:4, and between amino acids corresponding to amino acid numbers 536 and 542 of SEQ ID NO:4,

   wherein said protease recognition site is recognizable by a gut protease of western corn rootworm, and wherein said modified Cry3A toxin causes higher mortality to western corn rootworm than the mortality caused by said Cry3A toxin to western corn rootworm in an artificial diet bioassay.
2. The isolated nucleic acid molecule according to claim 1, wherein said gut protease is cathepsin G.
3. The isolated nucleic acid molecule according to claim 1, wherein said protease recognition site is located between amino acid numbers 107 and 115 of SEQ ID NO:4.
4. The isolated nucleic acid molecule according to claim 1, wherein said protease recognition site is located between amino acids corresponding to amino acid numbers 107 and 113 of SEQ.ID NO:4.
5. The isolated nucleic acid molecule according to claim 4, wherein said protease recognition site is located between amino acid numbers 107 and 113 of SEQ ID NO:4.
6. The isolated nucleic acid molecule according to claim 1, wherein said protease recognition site is located between amino acids corresponding to amino acid numbers 107 and 111 of SEQ ID NO:4.
7. The isolated nucleic acid molecule according to claim 6, wherein said protease recognition site is located between amino acid numbers 107 and 111 of SEQ ID NO:4.

8. The isolated nucleic acid molecule according to claim 1, wherein said protease site is located between amino acid numbers 536 and 542 of SEQ ID NO:4.

9. The isolated nucleic acid molecule according to claim 1, wherein said protease recognition site is located between amino acids corresponding to amino acid numbers 536 and 541 of SEQ ID NO:4.

10. The isolated nucleic acid molecule according to claim 9, wherein said additional protease site is located between amino acid numbers 536 and 541 of SEQ ID NO:4.

11. The isolated nucleic acid molecule according to claim 1, wherein said protease recognition site is located between amino acids corresponding to amino acid numbers 540 and 541 of SEQ ID NO:4.

12. The isolated nucleic acid molecule according to claim 11, wherein said protease site is located between amino acid numbers 540 and 541 of SEQ ID NO:4.

13. The isolated nucleic acid molecule according to claim 1, wherein said protease recognition site is located between amino acid numbers 107 and 115 of SEQ ID NO:4 and between amino acid numbers 536 and 542 of SEQ ID NO:4.

14. The isolated nucleic acid molecule according to claim 1, wherein said protease recognition site is located between amino acids corresponding to amino acid numbers 107 and 113 of SEQ ID NO:4 and between amino acids corresponding to amino acid numbers 540 and 541 of SEQ ID NO:4.

15. The isolated nucleic acid molecule according to claim 14, wherein said protease recognition site is located between amino acid numbers 107 and 113 of SEQ ID NO:4 and between amino acid numbers 540 and 541 of SEQ ID NO:4.

16. The isolated nucleic acid molecule according to claim 1, wherein said protease recognition site is located between amino acids corresponding to amino acid numbers. 107 and 111 of SEQ ID NO:4 and between amino acids corresponding to amino acid numbers 536 and 541 of SEQ ID NO:4.

17. The isolated nucleic acid molecule according to claim 16, wherein said protease recognition site is located between amino acid numbers 107 and 111 of SEQ ID NO:4 and between amino acid numbers 536 and 541 of SEQ ID NO:4.

18. The isolated nucleic acid molecule according to claim 1, wherein said protease recognition site is located between amino acids corresponding to amino acid numbers 107 and 111 of SEQ ID NO:4 and between amino acids corresponding to amino acid numbers 540 and 541 of SEQ ID NO:4.

19. The isolated nucleic acid molecule according to claim 18, wherein said protease recognition site is located between amino acid numbers 107 and 111 of SEQ ID NO:4 and between amino acid numbers 540 and 541 of SEQ ID NO:4.

20. The isolated nucleic acid molecule according to claim 1, wherein said modified Cry3A toxin causes at least 50% mortality to western corn rootworm to which said Cry3A toxin causes up to 30% mortality.

21. The isolated nucleic acid molecule according to claim 1, wherein said nucleotide sequence comprises nucleotides 1-1791 of SEQ ID NO: 6, nucleotides 1-1806 of SEQ ID NO: 8, nucleotides 1-1812 of SEQ ID NO: 10, nucleotides 1-1794 of SEQ ID NO: 12, nucleotides 1-1818 of SEQ ID NO: 14, nucleotides 1-1812 of SEQ ID NO: 16, nucleotides 1-1791 of SEQ ID NO: 18, or nucleotides 1-1818 of SEQ ID NO: 20.

22. The isolated nucleic acid molecule according to claim 1, wherein said modified Cry3A toxin comprises the amino acid sequence set forth in SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, or SEQ ID NO: 21.

23. The isolated nucleic acid molecule according to claim 1, wherein said modified Cry3A toxin is active against northern corn rootworm.

24. A chimeric construct comprising a heterologous promoter sequence operatively linked to the nucleic acid molecule of claim 1.

25. A recombinant vector comprising the chimeric construct of claim 24.

26. A transgenic non-human host cell comprising the chimeric construct of claim 24.

27. The transgenic host cell according to claim 26, which is a bacterial cell.

28. The transgenic host cell according to claim 26, which is a plant cell.

29. A transgenic plant comprising the transgenic plant cell of claim 28.

30. The transgenic plant according to claim 29, wherein said plant is a maize plant.

31. Transgenic seed from the transgenic plant of claim 29.

32. Transgenic seed from the maize plant of claim 30.

60
61

33. An isolated toxin produced by the expression of the nucleic acid molecule according to claim 1.

34. The isolated toxin according to claim 33, wherein said toxin is produced by the expression of a nucleic acid molecule comprising nucleotides 1-1791 of SEQ ID NO: 6, nucleotides 1-1806 of SEQ ID NO: 8, nucleotides 1-1812 of SEQ ID NO: 10, nucleotides 1-1794 of SEQ ID NO: 12, nucleotides 1-1818 of SEQ ID NO: 14, nucleotides 1-1812 of SEQ ID NO: 16, nucleotides 1-1791 of SEQ ID NO: 18, or nucleotides 1-1818 of SEQ ID NO: 20.

35. The isolated toxin according to claim 33, wherein said toxin has activity against northern corn rootworm.

36. The isolated toxin according to claim 33, wherein said toxin comprises the amino acid sequence set forth in SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, or SEQ ID NO: 21.

37. A composition comprising an effective amount of the toxin of claim 33 to cause mortality to western corn rootworm.

38. A transgenic maize plant comprising a nucleotide sequence that encodes a modified Cry3A toxin comprising a non-naturally occurring protease recognition site, wherein said protease recognition site modifies a Cry3A toxin and is located at a position selected from the group consisting of:

a) a position between amino acids corresponding to amino acid numbers 107 and 115 of SEQ ID NO:4;

b) a position between amino acids corresponding to amino acid numbers 536 and 542 of SEQ ID NO:4; and

c) a position between amino acids corresponding to amino acid numbers 107 and 115 of SEQ ID NO:4, and between amino acids corresponding to amino acid numbers 536 and 542 of SEQ ID NO:4,

wherein said protease recognition site is recognizable by a gut protease of western corn rootworm, and wherein said transgenic plant expresses said modified Cry3A toxin in root tissue at a level that causes mortality to western corn rootworm.

39. The transgenic maize plant according to claim 38, wherein said protease recognition site is located between amino acid numbers 107 and 111 of SEQ ID NO:4.

40. The transgenic maize plant according to claim 38, wherein said protease recognition site is located between amino acid numbers 540 and 541 of SEQ ID NO:4.

41. The transgenic maize plant according to claim 38, wherein said protease recognition site is located between amino acid numbers 107 and 111 of SEQ ID NO:4 and between amino acid numbers 540 and 541 of SEQ ID NO:4.

42. The transgenic maize plant according to claim 38, wherein said nucleotide sequence comprises nucleotides 1-1806 of SEQ ID NO: 8, nucleotides 1-1818 of SEQ ID NO: 14, or nucleotides 1-1791 of SEQ ID NO: 18.

43. The transgenic maize plant according to claim 38, wherein said modified Cry3A toxin comprises the amino acid sequence set forth in SEQ ID NO: 9, SEQ ID NO: 15, or SEQ ID NO: 19.

44. The transgenic maize plant according to claim 38, wherein said root tissue causes 100% mortality to western corn rootworm.

45. The transgenic maize plant according to claim 38, wherein said root tissue causes 90% mortality to western corn rootworm.

46. The transgenic maize plant according to claim 38, wherein said root tissue causes 80% mortality to western corn rootworm.

47. The transgenic maize plant according to claim 38, wherein said root tissue causes 70% mortality to western corn rootworm.

48. The transgenic maize plant according to claim 38, wherein said root tissue causes 60% mortality to western corn rootworm.

49. The transgenic maize plant according to claim 38, wherein said root tissue causes 50% mortality to western corn rootworm.

50. The transgenic maize plant according to claim 38, wherein said root tissue causes 40% mortality to western corn rootworm.

51. The transgenic maize plant according to claim 38, wherein said transgenic plant expresses said modified Cry3A toxin at a level sufficient to prevent western corn rootworm from severely pruning the roots of the transgenic plant.

52. The transgenic maize plant according to claim 38, wherein said transgenic plant expresses said modified Cry3A toxin at a level sufficient to prevent western corn rootworm feeding damage from causing the plant to lodge.

53. The transgenic maize plant according to any one of claims 38-52, which is an inbred plant.

54. The transgenic maize plant according to any one of claims 38-52, which is a hybrid plant.

55. Transgenic seed from the plant of claim 53.

56. Transgenic seed from the plant of claim 54.

57. A modified Cry3A toxin comprising a non-naturally occurring protease recognition site, wherein said protease recognition site modifies a Cry3A toxin and is located at a position selected from the group consisting of:

a) a position between amino acids corresponding to amino acid numbers 107 and 115 of SEQ ID NO:4;

b) a position between amino acids corresponding to amino acid numbers 536 and 542 of SEQ ID NO:4; and

c) a position between amino acids corresponding to amino acid numbers 107 and 115 of SEQ ID NO:4, and between amino acids corresponding to amino acid numbers 536 and 542 of SEQ ID NO:4,

wherein said protease recognition site is recognizable by a gut protease of western corn rootworm, and wherein said modified Cry3A toxin causes higher mortality to western corn rootworm than the mortality caused by said Cry3A toxin to western corn rootworm in an artificial diet bioassay.

58. The modified Cry3A toxin according to claim 57, wherein said gut protease is cathepsin G.

59. The modified Cry3A toxin according to claim 57, wherein said protease recognition site is located between amino acid numbers 107 and 115 of SEQ ID NO:4.

60. The modified Cry3A toxin according to claim 57, wherein said protease recognition site is located between amino acids corresponding to amino acid numbers 107 and 113 of SEQ ID NO:4.

61. The modified Cry3A toxin according to claim 60, wherein said protease recognition site is located between amino acid numbers 107 and 113 of SEQ ID NO:4.

62. The modified Cry3A toxin according to claim 57, wherein said protease recognition site is located between amino acids corresponding to amino acid numbers 107 and 111 of SEQ ID NO:4.

63. The modified Cry3A toxin according to claim 62, wherein said protease recognition site is located between amino acid numbers 107 and 111 of SEQ ID NO:4.

64. The modified Cry3A toxin according to claim 57, wherein said protease site is located between amino acid numbers 536 and 542 of SEQ ID NO:4.

65. The modified Cry3A toxin according to claim 57, wherein said protease recognition site is located between amino acids corresponding to amino acid numbers 536 and 541 of SEQ ID NO:4.

66. The modified Cry3A toxin according to claim 65, wherein said additional protease site is located between amino acid numbers 536 and 541 of SEQ ID NO:4.

67. The modified Cry3A toxin according to claim 57, wherein said protease recognition site is located between amino acids corresponding to amino acid numbers 540 and 541 of SEQ ID NO:4.

68. The modified Cry3A toxin according to claim 67, wherein said protease site is located between amino acid numbers 540 and 541 of SEQ ID NO:4.

69. The modified Cry3A toxin according to claim 57, wherein said protease recognition site is located between amino acid numbers 107 and 115 and between amino acid numbers 536 and 542 of SEQ ID NO:4.

70. The modified Cry3A toxin according to claim 57, wherein said protease recognition site is located between amino acids corresponding to amino acid numbers 107 and 113 of SEQ ID NO:4 and between amino acids corresponding to amino acid numbers 540 and 541 of SEQ ID NO:4.

71. The modified Cry3A toxin according to claim 70, wherein said protease recognition site is located between amino acid numbers 107 and 113 of SEQ ID NO:4 and between amino acid numbers 541 and 541 of SEQ ID NO:4.

72. The modified Cry3A toxin according to claim 57, wherein said protease recognition site is located between amino acids corresponding to amino acid numbers 107 and 111 of SEQ ID NO:4 and between amino acids corresponding to amino acid numbers 536 and 541 of SEQ ID NO:4.

73. The modified Cry3A toxin according to claim 72, wherein said protease recognition site is located between amino acid numbers 107 and 111 of SEQ ID NO:4 and between amino acid numbers 536 and 541 of SEQ ID NO:4.

74. The modified Cry3A toxin according to claim 57, wherein said protease recognition site is located between amino acids corresponding to amino acid numbers 107 and 111 of SEQ ID NO:4 and between amino acids corresponding to amino acid numbers 540 and 541 of SEQ ID NO:4.

75. The modified Cry3A toxin according to claim 74, wherein said protease recognition site is located between amino acid numbers 107 and 111 of SEQ ID NO:4 and between amino acid numbers 540 and 541 of SEQ ID NO:4.

76. The modified Cry3A toxin according to claim 57, wherein said modified Cry3A toxin causes at least 50% mortality to western corn rootworm to which said Cry3A toxin causes up to 30% mortality.

77. The modified Cry3A toxin according to claim 57 which is encoded by nucleotides 1-1791 of SEQ ID NO: 6, nucleotides 1-1806 of SEQ ID NO: 8, nucleotides 1-1812 of SEQ ID NO: 10, nucleotides 1-1794 of SEQ ID NO: 12, nucleotides 1-1818 of SEQ ID NO: 14, nucleotides 1-1812 of SEQ ID NO: 16, nucleotides 1-1791 of SEQ ID NO: 18, or nucleotides 1-1818 of SEQ ID NO: 20.

78. The modified Cry3A toxin according to claim 57 comprising the amino acid sequence set forth in SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, or SEQ ID NO: 21.

79. The modified Cry3A toxin according to claim 57 which is active against northern corn rootworm.

80. A method of controlling infestation of maize plants by western corn rootworm, the method comprising:

(a) providing the transgenic maize plant according to claim 38; and

(b) contacting said western corn rootworm with the plant.

81. A method of producing a modified Cry3A toxin, comprising:
(a) obtaining the transgenic host cell according to claim 27;
(b) culturing the transgenic host cell under conditions that allow the expression of the modified Cry3A toxin; and
(c) recovering the expressed modified Cry3A toxin.

82. A method of producing insect-resistant plants, comprising:
(a) stably integrating the nucleic acid molecule according to claim 1 into the genome of plant cells; and
(b) regenerating stably transformed plants from said transformed plant cells, wherein said stably transformed plants express an effective amount of a modified Cry3A toxin to render said transformed plant resistant to at least western corn rootworm.

83. A method of controlling at least western corn rootworm, comprising delivering orally to western corn rootworm an effective amount of a toxin according to claim 33.

84. A method of making a modified Cry3A toxin, comprising:
(a) obtaining a *cry*3A gene which encodes a Cry3A toxin;
(b) identifying a gut protease of western corn rootworm;
(c) obtaining a nucleotide sequence which encodes a recognition site for said gut protease;
(d) inserting said nucleotide sequence into said *cry*3A gene, such that said recognition site is located in said Cry3A toxin at a position between amino acids corresponding to amino acid numbers 107 and 115 of SEQ ID NO:4, or at a position between amino acids corresponding to amino acid numbers 536 and 542 of SEQ ID NO:4, or at a position between amino acids corresponding to amino acid numbers 107 and 115 of SEQ ID NO:4 and between amino acids corresponding to amino acid numbers 536 and 542 of SEQ ID NO:4, thus creating a modified *cry*3A gene;
(e) inserting said modified *cry*3A gene into an expression cassette; and
(f) transforming said expression cassette into a non-human host cell, wherein said host cell produces a modified Cry3A toxin.

85. A modified *cry*3A gene comprising a nucleotide sequence that encodes a modified Cry3A toxin comprising a non-naturally occurring protease recognition site, wherein said modified *cry*3A gene comprises a coding sequence encoding said protease

recognition site, wherein said coding sequence modifies a *cry*3A gene and is inserted at a position selected from the group consisting of:

a) a position between the codons that code for amino acids corresponding to amino acid numbers 107 and 115 of SEQ ID NO:4;

b) a position between the codons that code for amino acids corresponding to amino acid numbers 536 and 542 of SEQ ID NO:4; and

c) a position between the codons that code for amino acids corresponding to amino acid numbers 107 and 115 of SEQ ID NO:4, and between codons that code for amino acids corresponding to amino acid numbers 536 and 542 of SEQ ID NO:4,

wherein said protease recognition site is recognizable by a gut protease of western corn rootworm, and wherein said modified Cry3A toxin causes higher mortality to western corn rootworm than the mortality caused by said Cry3A toxin to western corn rootworm in an artificial diet bioassay.

86. The modified *cry*3A gene according to claim 85, wherein said gut protease is cathepsin G.

87. The modified *cry*3A gene according to claim 85, wherein said nucleotide sequence comprises nucleotides 1-1791 of SEQ ID NO: 6, nucleotides 1-1806 of SEQ ID NO: 8, nucleotides 1-1812 of SEQ ID NO: 10, nucleotides 1-1794 of SEQ ID NO: 12, nucleotides 1-1818 of SEQ ID NO: 14, nucleotides 1-1812 of SEQ ID NO: 16, nucleotides 1-1791 of SEQ ID NO: 18, or nucleotides 1-1818 of SEQ ID NO: 20.

88. The modified *cry*3A gene according to claim 85, wherein said modified Cry3A toxin comprises the amino acid sequence set forth in SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, or SEQ ID NO: 21.

89. The modified *cry*3A gene according to claim 85, wherein said modified Cry3A toxin is active against northern corn rootworm.

90. A chimeric construct comprising a heterologous promoter sequence operatively linked to the modified *cry*3A gene of claim 85.

91. A recombinant vector comprising the chimeric construct of claim 90.

92. A transgenic non-human host cell comprising the chimeric construct of claim 90.

93. The transgenic host cell according to claim 92, which is a bacterial cell.

94. The transgenic host cell according to claim 92, which is a plant cell.

95. A transgenic plant comprising the transgenic plant cell of claim 94.

96. The transgenic plant according to claim 95, wherein said plant is a maize plant.

97. Transgenic seed from the transgenic plant of claim 95.

98. Transgenic seed from the maize plant of claim 96.

99. A transgenic maize plant comprising a modified *cry*3A gene comprising a nucleotide sequence that encodes a modified Cry3A toxin comprising a non-naturally occurring protease recognition site, wherein said modified *cry*3A gene comprises a coding sequence encoding said protease recognition site, wherein said coding sequence modifies a *cry*3A gene and is inserted at a position selected from the group consisting of:

a) a position between the codons that code for amino acids corresponding to amino acid numbers 107 and 115 of SEQ ID NO:4;

b) a position between the codons that code for amino acids corresponding to amino acid numbers 536 and 542 of SEQ ID NO:4; and

c) a position between the codons that code for amino acids corresponding to amino acid numbers 107 and 115 of SEQ ID NO:4, and between codons that code for amino acids corresponding to amino acid numbers 536 and 542 of SEQ ID NO:4,

wherein said protease recognition site is recognizable by a gut protease of western corn rootworm, and wherein said transgenic plant expresses said modified Cry3A toxin in root tissue at a level that causes mortality to western corn rootworm.

100. The transgenic maize plant according to claim 99, wherein said nucleotide sequence comprises nucleotides 1-1806 of SEQ ID NO: 8, nucleotides 1-1818 of SEQ ID NO: 14, or nucleotides 1-1791 of SEQ ID NO: 18.

101. The transgenic maize plant according to claim 99, wherein said modified Cry3A toxin comprises the amino acid sequence set forth in SEQ ID NO: 9, SEQ ID NO: 15, or SEQ ID NO: 19.

102. The transgenic maize plant according to any one of claims 99-101, which is an inbred plant.

103. The transgenic maize plant according to any one of claims 99-101, which is a hybrid plant.

104. Transgenic seed from the plant of claim 102.

105. Transgenic seed from the plant of claim 103.

SEQUENCE LISTING

<110> Syngenta Participations AG

<120> Modified Cry3A Toxins and Nucleic Acid Sequences Coding Therefor

<130> 60065/PCT

<140>
<141>

<150> US 60/316421
<151> 2001-08-31

<160> 34

<170> PatentIn Ver. 3.0

<210> 1
<211> 1932
<212> DNA
<213> Bacillus thuringiensis

<220>
<221> CDS
<222> (1)..(1932)
<223> Native cry3A coding sequence according to Sekar et al. 1987, Proc. Natl. Acad. Sci. 84:7036-7040

<400> 1

```
atg aat ccg aac aat cga agt gaa cat gat aca ata aaa act act gaa       48
Met Asn Pro Asn Asn Arg Ser Glu His Asp Thr Ile Lys Thr Thr Glu
1               5                  10                  15

aat aat gag gtg cca act aac cat gtt caa tat cct tta gcg gaa act       96
Asn Asn Glu Val Pro Thr Asn His Val Gln Tyr Pro Leu Ala Glu Thr
            20                  25                  30

cca aat cca aca cta gaa gat tta aat tat aaa gag ttt tta aga atg      144
Pro Asn Pro Thr Leu Glu Asp Leu Asn Tyr Lys Glu Phe Leu Arg Met
        35                  40                  45

act gca gat aat aat acg gaa gca cta gat agc tct aca aca aaa gat      192
Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys Asp
    50                  55                  60

gtc att caa aaa ggc att tcc gta gta ggt gat ctc cta ggc gta gta      240
Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val Val
65                  70                  75                  80

ggt ttc ccg ttt ggt gga gcg ctt gtt tcg ttt tat aca aac ttt tta      288
Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe Leu
                85                  90                  95
```

```
aat act att tgg cca agt gaa gac ccg tgg aag gct ttt atg gaa caa      336
Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu Gln
            100                 105                 110

gta gaa gca ttg atg gat cag aaa ata gct gat tat gca aaa aat aaa      384
Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn Lys
        115                 120                 125

gct ctt gca gag tta cag ggc ctt caa aat aat gtc gaa gat tat gtg      432
Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr Val
    130                 135                 140

agt gca ttg agt tca tgg caa aaa aat cct gtg agt tca cga aat cca      480
Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn Pro
145                 150                 155                 160

cat agc cag ggg cgg ata aga gag ctg ttt tct caa gca gaa agt cat      528
His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser His
                165                 170                 175

ttt cgt aat tca atg cct tcg ttt gca att tct gga tac gag gtt cta      576
Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val Leu
            180                 185                 190

ttt cta aca aca tat gca caa gct gcc aac aca cat tta ttt tta cta      624
Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu Leu
        195                 200                 205

aaa gac gct caa att tat gga gaa gaa tgg gga tac gaa aaa gaa gat      672
Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu Asp
    210                 215                 220

att gct gaa ttt tat aaa aga caa cta aaa ctt acg caa gaa tat act      720
Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr Thr
225                 230                 235                 240

gac cat tgt gtc aaa tgg tat aat gtt gga tta gat aaa tta aga ggt      768
Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg Gly
                245                 250                 255

tca tct tat gaa tct tgg gta aac ttt aac cgt tat cgc aga gag atg      816
Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu Met
            260                 265                 270

aca tta aca gta tta gat tta att gca cta ttt cca ttg tat gat gtt      864
Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp Val
        275                 280                 285

cgg cta tac cca aaa gaa gtt aaa acc gaa tta aca aga gac gtt tta      912
Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val Leu
    290                 295                 300

aca gat cca att gtc gga gtc aac aac ctt agg ggc tat gga aca acc      960
Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr Thr
```

```
305                 310                 315                 320

ttc tct aat ata gaa aat tat att cga aaa cca cat cta ttt gac tat     1008
Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp Tyr
                325                 330                 335

ctg cat aga att caa ttt cac acg cgg ttc caa cca gga tat tat gga     1056
Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr Gly
            340                 345                 350

aat gac tct ttc aat tat tgg tcc ggt aat tat gtt tca act aga cca     1104
Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg Pro
        355                 360                 365

agc ata gga tca aat gat ata atc aca tct cca ttc tat gga aat aaa     1152
Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn Lys
    370                 375                 380

tcc agt gaa cct gta caa aat tta gaa ttt aat gga gaa aaa gtc tat     1200
Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val Tyr
385                 390                 395                 400

aga gcc gta gca aat aca aat ctt gcg gtc tgg ccg tcc gct gta tat     1248
Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val Tyr
                405                 410                 415

tca ggt gtt aca aaa gtg gaa ttt agc caa tat aat gat caa aca gat     1296
Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr Asp
            420                 425                 430

gaa gca agt aca caa acg tac gac tca aaa aga aat gtt ggc gcg gtc     1344
Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala Val
        435                 440                 445

agc tgg gat tct atc gat caa ttg cct cca gaa aca aca gat gaa cct     1392
Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu Pro
    450                 455                 460

cta gaa aag gga tat agc cat caa ctc aat tat gta atg tgc ttt tta     1440
Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe Leu
465                 470                 475                 480

atg cag ggt agt aga gga aca atc cca gtg tta act tgg aca cat aaa     1488
Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His Lys
                485                 490                 495

agt gta gac ttt ttt aac atg att gat tcg aaa aaa att aca caa ctt     1536
Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln Leu
            500                 505                 510

ccg tta gta aag gca tat aag tta caa tct ggt gct tcc gtt gtc gca     1584
Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val Ala
        515                 520                 525
```

```
ggt cct agg ttt aca gga gga gat atc att caa tgc aca gaa aat gga     1632
Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn Gly
    530                 535                 540

agt gcg gca act att tac gtt aca ccg gat gtg tcg tac tct caa aaa     1680
Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln Lys
545                 550                 555                 560

tat cga gct aga att cat tat gct tct aca tct cag ata aca ttt aca     1728
Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe Thr
                565                 570                 575

ctc agt tta gac ggg gca cca ttt aat caa tac tat ttc gat aaa acg     1776
Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp Lys Thr
            580                 585                 590

ata aat aaa gga gac aca tta acg tat aat tca ttt aat tta gca agt     1824
Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala Ser
        595                 600                 605

ttc agc aca cca ttc gaa tta tca ggg aat aac tta caa ata ggc gtc     1872
Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly Val
    610                 615                 620

aca gga tta agt gct gga gat aaa gtt tat ata gac aaa att gaa ttt     1920
Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu Phe
625                 630                 635                 640

att cca gtg aat                                                    1932
Ile Pro Val Asn
```

<210> 2
<211> 644
<212> PRT
<213> Bacillus thuringiensis

<400> 2

```
Met Asn Pro Asn Asn Arg Ser Glu His Asp Thr Ile Lys Thr Thr Glu
1               5                   10                  15

Asn Asn Glu Val Pro Thr Asn His Val Gln Tyr Pro Leu Ala Glu Thr
            20                  25                  30

Pro Asn Pro Thr Leu Glu Asp Leu Asn Tyr Lys Glu Phe Leu Arg Met
        35                  40                  45

Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys Asp
    50                  55                  60

Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val Val
65                  70                  75                  80

Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe Leu
```

85 90 95

Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu Gln
100 105 110

Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn Lys
115 120 125

Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr Val
130 135 140

Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn Pro
145 150 155 160

His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser His
165 170 175

Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val Leu
180 185 190

Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu Leu
195 200 205

Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu Asp
210 215 220

Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr Thr
225 230 235 240

Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg Gly
245 250 255

Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu Met
260 265 270

Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp Val
275 280 285

Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val Leu
290 295 300

Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr Thr
305 310 315 320

Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp Tyr
325 330 335

Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr Gly
340 345 350

Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg Pro
355 360 365

Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn Lys

370 375 380

Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val Tyr
385 390 395 400

Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val Tyr
405 410 415

Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr Asp
420 425 430

Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala Val
435 440 445

Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu Pro
450 455 460

Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe Leu
465 470 475 480

Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His Lys
485 490 495

Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln Leu
500 505 510

Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val Ala
515 520 525

Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn Gly
530 535 540

Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln Lys
545 550 555 560

Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe Thr
565 570 575

Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp Lys Thr
580 585 590

Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala Ser
595 600 605

Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly Val
610 615 620

Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu Phe
625 630 635 640

Ile Pro Val Asn

<210> 3
<211> 1803
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (1)..(1794)
<223> Maize optimized cry3A coding sequence

<400> 3

```
atg acg gcc gac aac aac acc gag gcc ctg gac agc agc acc acc aag      48
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1               5                  10                  15

gac gtg atc cag aag ggc atc agc gtg gtg ggc gac ctg ctg ggc gtg      96
Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
            20                  25                  30

gtg ggc ttc ccc ttc ggc ggc gcc ctg gtg agc ttc tac acc aac ttc     144
Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
        35                  40                  45

ctg aac acc atc tgg ccc agc gag gac ccc tgg aag gcc ttc atg gag     192
Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
    50                  55                  60

cag gtg gag gcc ctg atg gac cag aag atc gcc gac tac gcc aag aac     240
Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65                  70                  75                  80

aag gca ctg gcc gag cta cag ggc ctc cag aac aac gtg gag gac tat     288
Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
                85                  90                  95

gtg agc gcc ctg agc agc tgg cag aag aac ccc gtc tcg agc cgc aac     336
Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn
            100                 105                 110

ccc cac agc cag ggc cgc atc cgc gag ctg ttc agc cag gcc gag agc     384
Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser
        115                 120                 125

cac ttc cgc aac agc atg ccc agc ttc gcc atc agc ggc tac gag gtg     432
His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val
    130                 135                 140

ctg ttc ctg acc acc tac gcc cag gcc gcc aac acc cac ctg ttc ctg     480
Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu
145                 150                 155                 160

ctg aag gac gcc caa atc tac gga gag gag tgg ggc tac gag aag gag     528
Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu
                165                 170                 175
```

```
gac atc gcc gag ttc tac aag cgc cag ctg aag ctg acc cag gag tac      576
Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr
            180                 185                 190

acc gac cac tgc gtg aag tgg tac aac gtg ggt cta gac aag ctc cgc      624
Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg
        195                 200                 205

ggc agc agc tac gag agc tgg gtg aac ttc aac cgc tac cgc cgc gag      672
Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu
    210                 215                 220

atg acc ctg acc gtg ctg gac ctg atc gcc ctg ttc ccc ctg tac gac      720
Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp
225                 230                 235                 240

gtg cgc ctg tac ccc aag gag gtg aag acc gag ctg acc cgc gac gtg      768
Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val
                245                 250                 255

ctg acc gac ccc atc gtg ggc gtg aac aac ctg cgc ggc tac ggc acc      816
Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr
            260                 265                 270

acc ttc agc aac atc gag aac tac atc cgc aag ccc cac ctg ttc gac      864
Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp
        275                 280                 285

tac ctg cac cgc atc cag ttc cac acg cgt ttc cag ccc ggc tac tac      912
Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr
    290                 295                 300

ggc aac gac agc ttc aac tac tgg agc ggc aac tac gtg agc acc cgc      960
Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg
305                 310                 315                 320

ccc agc atc ggc agc aac gac atc atc acc agc ccc ttc tac ggc aac     1008
Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn
                325                 330                 335

aag agc agc gag ccc gtg cag aac ctt gag ttc aac ggc gag aag gtg     1056
Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val
            340                 345                 350

tac cgc gcc gtg gct aac acc aac ctg gcc gtg tgg ccc tct gca gtg     1104
Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val
        355                 360                 365

tac agc ggc gtg acc aag gtg gag ttc agc cag tac aac gac cag acc     1152
Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr
    370                 375                 380

gac gag gcc agc acc cag acc tac gac agc aag cgc aac gtg ggc gcc     1200
```

```
Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala
385                 390                 395                 400

gtg agc tgg gac agc atc gac cag ctg ccc ccc gag acc acc gac gag    1248
Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu
                405                 410                 415

ccc ctg gag aag ggc tac agc cac cag ctg aac tac gtg atg tgc ttc    1296
Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe
            420                 425                 430

ctg atg cag ggc agc cgc ggc acc atc ccc gtg ctg acc tgg acc cac    1344
Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His
        435                 440                 445

aag agc gtc gac ttc ttc aac atg atc gac agc aag aag atc acc cag    1392
Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln
    450                 455                 460

ctg ccc ctg gtg aag gcc tac aag ctc cag agc ggc gcc agc gtg gtg    1440
Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val
465                 470                 475                 480

gca ggc ccc cgc ttc acc ggc ggc gac atc atc cag tgc acc gag aac    1488
Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn
                485                 490                 495

ggc agc gcc gcc acc atc tac gtg acc ccc gac gtg agc tac agc cag    1536
Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln
            500                 505                 510

aag tac cgc gcc cgc atc cac tac gcc agc acc agc cag atc acc ttc    1584
Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe
        515                 520                 525

acc ctg agc ctg gac ggg gcc ccc ttc aac caa tac tac ttc gac aag    1632
Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp Lys
    530                 535                 540

acc atc aac aag ggc gac acc ctg acc tac aac agc ttc aac ctg gcc    1680
Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala
545                 550                 555                 560

agc ttc agc acc cct ttc gag ctg agc ggc aac aac ctc cag atc ggc    1728
Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly
                565                 570                 575

gtg acc ggc ctg agc gcc ggc gac aag gtg tac atc gac aag atc gag    1776
Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu
            580                 585                 590

ttc atc ccc gtg aac tag atctgagct                                  1803
Phe Ile Pro Val Asn
        595
```

<210> 4
<211> 597
<212> PRT
<213> Cry3A encoded by SEQ ID NO:3.

<400> 4

```
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1               5                  10                  15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
            20                  25                  30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
        35                  40                  45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
    50                  55                  60

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

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
                85                  90                  95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn
            100                 105                 110

Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser
        115                 120                 125

His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val
    130                 135                 140

Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu
145                 150                 155                 160

Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu
                165                 170                 175

Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr
            180                 185                 190

Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg
        195                 200                 205

Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu
    210                 215                 220

Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp
225                 230                 235                 240

Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val
```

245 250 255

Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr
260 265 270

Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp
275 280 285

Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr
290 295 300

Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg
305 310 315 320

Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn
325 330 335

Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val
340 345 350

Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val
355 360 365

Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr
370 375 380

Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala
385 390 395 400

Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu
405 410 415

Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe
420 425 430

Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His
435 440 445

Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln
450 455 460

Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val
465 470 475 480

Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn
485 490 495

Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln
500 505 510

Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe
515 520 525

Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp Lys

530 535 540

Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala
545 550 555 560

Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly
565 570 575

Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu
580 585 590

Phe Ile Pro Val Asn
595

<210> 5
<211> 7208
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<223> pCIB6850

<400> 5
gatccaccat gacggccgac aacaacaccg aggccctgga cagcagcacc accaaggacg 60
tgatccagaa gggcatcagc gtggtgggcg acctgctggg cgtggtgggc ttcccccttcg 120
gcggcgccct ggtgagcttc tacaccaact tcctgaacac catctggccc agcgaggacc 180
cctggaaggc cttcatggag caggtggagg ccctgatgga ccagaagatc gccgactacg 240
ccaagaacaa ggcactggcc gagctacagg gcctccagaa caacgtggag gactatgtga 300
gcgccctgag cagctggcag aagaaccccg tctcgagccg caacccccac agccagggcc 360
gcatccgcga gctgttcagc caggccgaga gccacttccg caacagcatg cccagcttcg 420
ccatcagcgg ctacgaggtg ctgttcctga ccacctacgc ccaggccgcc aacacccacc 480
tgttcctgct gaaggacgcc caaatctacg gagaggagtg gggctacgag aaggaggaca 540
tcgccgagtt ctacaagcgc cagctgaagc tgacccagga gtacaccgac cactgcgtga 600
agtggtacaa cgtgggtcta gacaagctcc gcggcagcag ctacgagagc tgggtgaact 660
tcaaccgcta ccgccgcgag atgaccctga ccgtgctgga cctgatcgcc ctgttccccc 720
tgtacgacgt gcgcctgtac cccaaggagg tgaagaccga gctgacccgc gacgtgctga 780
ccgaccccat cgtgggcgtg aacaacctgc gcggctacgg caccaccttc agcaacatcg 840
agaactacat ccgcaagccc cacctgttcg actacctgca ccgcatccag ttccacacgc 900
gtttccagcc cggctactac ggcaacgaca gcttcaacta ctggagcggc aactacgtga 960
gcacccgccc cagcatcggc agcaacgaca tcatcaccag ccccttctac ggcaacaaga 1020
gcagcgagcc cgtgcagaac cttgagttca acggcgagaa ggtgtaccgc gccgtggcta 1080
acaccaacct ggccgtgtgg ccctctgcag tgtacagcgg cgtgaccaag gtggagttca 1140
gccagtacaa cgaccagacc gacgaggcca gcacccagac ctacgacagc aagcgcaacg 1200
tgggcgccgt gagctgggac agcatcgacc agctgccccc cgagaccacc gacgagcccc 1260
tggagaaggg ctacagccac cagctgaact acgtgatgtg cttcctgatg cagggcagcc 1320
gcggcaccat ccccgtgctg acctggaccc acaagagcgt cgacttcttc aacatgatcg 1380
acagcaagaa gatcacccag ctgcccctgg tgaaggccta caagctccag agcggcgcca 1440
gcgtggtggc aggccccccgc ttcaccggcg gcgacatcat ccagtgcacc gagaacggca 1500
gcgccgccac catctacgtg accccgacg tgagctacag ccagaagtac cgcgcccgca 1560
tccactacgc cagcaccagc cagatcacct tcaccctgag cctggacggg gccccccttca 1620
accaatacta cttcgacaag accatcaaca agggcgacac cctgacctac aacagcttca 1680
acctggccag cttcagcacc cctttcgagc tgagcggcaa caacctccag atcggcgtga 1740

ccggcctgag cgccggcgac aaggtgtaca tcgacaagat cgagttcatc cccgtgaact 1800
agatctgagc tcaagatctg ttgtacaaaa accagcaact cactgcactg cacttcactt 1860
cacttcactg tatgaataaa agtctggtgt ctggttcctg atcgatgact gactactcca 1920
ctttgtgcag aacttagtat gtatttgtat ttgtaaaata cttctatcaa taaaatttct 1980
aattcctaaa accaaaatcc agtgggtacc gaattcactg gccgtcgttt tacaacgtcg 2040
tgactgggaa aaccctggcg ttacccaact taatcgcctt gcagcacatc ccccctttcgc 2100
cagctggcgt aatagcgaag aggcccgcac cgatcgccct tcccaacagt tgcgcagcct 2160
gaatggcgaa tggcgcctga tgcggtattt tctccttacg catctgtgcg gtatttcaca 2220
ccgcatatgg tgcactctca gtacaatctg ctctgatgcc gcatagttaa gccagccccg 2280
acacccgcca acacccgctg acgcgccctg acgggcttgt ctgctcccgg catccgctta 2340
cagacaagct gtgaccgtct ccgggagctg catgtgtcag aggttttcac cgtcatcacc 2400
gaaacgcgcg agacgaaagg gcctcgtgat acgcctattt ttataggtta atgtcatgat 2460
aataatggtt tcttagacgt caggtggcac ttttcgggga aatgtgcgcg gaacccctat 2520
ttgtttattt ttctaaatac attcaaatat gtatccgctc atgagacaat aaccctgata 2580
aatgcttcaa taatattgaa aaaggaagag tatgagtatt caacatttcc gtgtcgccct 2640
tattcccttt tttgcggcat tttgccttcc tgtttttgct cacccagaaa cgctggtgaa 2700
agtaaaagat gctgaagatc agttgggtgc acgagtgggt tacatcgaac tggatctcaa 2760
cagcggtaag atccttgaga gttttcgccc cgaagaacgt tttccaatga tgagcacttt 2820
taaagttctg ctatgtggcg cggtattatc ccgtattgac gccgggcaag agcaactcgg 2880
tcgccgcata cactattctc agaatgactt ggttgagtac tcaccagtca cagaaaagca 2940
tcttacggat ggcatgacag taagagaatt atgcagtgct gccataacca tgagtgataa 3000
cactgcggcc aacttacttc tgacaacgat cggaggaccg aaggagctaa ccgcttttt 3060
gcacaacatg ggggatcatg taactcgcct tgatcgttgg gaaccggagc tgaatgaagc 3120
cataccaaac gacgagcgtg acaccacgat gcctgtagca atggcaacaa cgttgcgcaa 3180
actattaact ggcgaactac ttactctagc ttcccggcaa caattaatag actggatgga 3240
ggcggataaa gttgcaggac cacttctgcg ctcggccctt ccggctggct ggtttattgc 3300
tgataaatct ggagccggtg agcgtgggtc tcgcggtatc attgcagcac tggggccaga 3360
tggtaagccc tcccgtatcg tagttatcta cacgacgggg agtcaggcaa ctatggatga 3420
acgaaataga cagatcgctg agataggtgc ctcactgatt aagcattggt aactgtcaga 3480
ccaagtttac tcatatatac tttagattga tttaaaactt cattttaat ttaaaaggat 3540
ctaggtgaag atcctttttg ataatctcat gaccaaaatc ccttaacgtg agttttcgtt 3600
ccactgagcg tcagaccccg tagaaaagat caaaggatct tcttgagatc ctttttttct 3660
gcgcgtaatc tgctgcttgc aaacaaaaaa accaccgcta ccagcggtgg tttgtttgcc 3720
ggatcaagag ctaccaactc tttttccgaa ggtaactggc ttcagcagag cgcagatacc 3780
aaatactgtc cttctagtgt agccgtagtt aggccaccac ttcaagaact ctgtagcacc 3840
gcctacatac ctcgctctgc taatcctgtt accagtggct gctgccagtg gcgataagtc 3900
gtgtcttacc gggttggact caagacgata gttaccggat aaggcgcagc ggtcgggctg 3960
aacgggggt tcgtgcacac agcccagctt ggagcgaacg acctacaccg aactgagata 4020
cctacagcgt gagctatgag aaagcgccac gcttcccgaa gggagaaagg cggacaggta 4080
tccggtaagc ggcagggtcg gaacaggaga gcgcacgagg gagcttccag ggggaaacgc 4140
ctggtatctt tatagtcctg tcgggtttcg ccacctctga cttgagcgtc gatttttgtg 4200
atgctcgtca ggggggcgga gcctatggaa aaacgccagc aacgcggcct ttttacggtt 4260
cctggccttt tgctggcctt ttgctcacat gttctttcct gcgttatccc ctgattctgt 4320
ggataaccgt attaccgcct ttgagtgagc tgataccgct cgccgcagcc gaacgaccga 4380
gcgcagcgag tcagtgagcg aggaagcgga agagcgccca atacgcaaac cgcctctccc 4440
cgcgcgttgg ccgattcatt aatgcagctg gcacgacagg tttcccgact ggaaagcggg 4500
cagtgagcgc aacgcaatta atgtgagtta gctcactcat taggcacccc aggctttaca 4560
ctttatgctt ccggctcgta tgttgtgtgg aattgtgagc ggataacaat ttcacacagg 4620
aaacagctat gaccatgatt acgccaagct tgcacatgac aacaattgta agaggatgga 4680
gaccacaacg atccaacaat acttctgcga cgggctgtga agtatagaga agttaaacgc 4740
ccaaaagcca ttgtgtttgg aatttttagt tattctattt ttcatgatgt atcttcctct 4800
aacatgcctt aatttgcaaa tttggtataa ctactgattg aaaatatatg tatgtaaaaa 4860
aatactaagc atatttgtga agctaaacat gatgttattt aagaaaatat gttgttaaca 4920
gaataagatt aatatcgaaa tggaaacatc tgtaaattag aatcatctta caagctaaga 4980

gatgttcacg ctttgagaaa cttcttcaga tcatgaccgt agaagtagct ctccaagact 5040
caacgaaggc tgctgcaatt ccacaaatgc atgacatgca tccttgtaac cgtcgtcgcc 5100
gctataaaca cggataactc aattcccctgc tccatcaatt tagaaatgag caagcaagca 5160
cccgatcgct caccccatat gcaccaatct gactcccaag tctctgtttc gcattagtac 5220
cgccagcact ccacctatag ctaccaattg agacctttcc agcctaagca gatcgattga 5280
tcgttagagt caaagagttg gtggtacggg tactttaact accatggaat gatggggcgt 5340
gatgtagagc ggaaagcgcc tccctacgcg gaacaacacc ctcgccatgc cgctcgacta 5400
cagcctcctc ctcgtcggcc gcccacaacg agggagcccg tggtcgcagc caccgaccag 5460
catgtctctg tgtcctcgtc cgacctcgac atgtcatggc aaacagtcgg acgccagcac 5520
cagactgacg acatgagtct ctgaagagcc cgccacctag aaagatccga gccctgctgc 5580
tggtagtggt aaccattttc gtcgcgctga cgcggagagc gagaggccag aaatttatag 5640
cgactgacgc tgtggcaggc acgctatcgg aggttacgac gtggcgggtc actcgacgcg 5700
gagttcacag gtcctatcct tgcatcgctc gggccggagt ttacgggact tatccttacg 5760
acgtgctcta aggttgcgat aacgggcgga ggaaggcgtg tggcgtgcgg agacggttta 5820
tacacgtagt gtgcgggagt gtgtttcgta gacgcgggaa agcacgacga cttacgaagg 5880
ttagtggagg aggaggacac actaaaatca ggacgcaaga aactcttcta ttatagtagt 5940
agagaagaga ttataggagt gtgggttgat tctaaagaaa atcgacgcag gacaaccgtc 6000
aaaacgggtg ctttaatata gtagatatat atatatagag agagagagaa agtacaaagg 6060
atgcatttgt gtctgcatat gatcggagta ttactaacgg ccgtcgtaag aaggtccatc 6120
atgcgtggag cgagcccatt tggttggttg tcaggccgca gttaaggcct ccatatatga 6180
ttgtcgtcgg gcccataaca gcatctcctc caccagttta ttgtaagaat aaattaagta 6240
gagatatttg tcgtcgggca gaagaaactt ggacaagaag aagaagcaag ctaggccaat 6300
ttcttgccgg caagaggaag atagtggcct ctagtttata tatcggcgtg atgatgatgc 6360
tcctagctag aaatgagaga agaaaaacgg acgcgtgttt ggtgtgtgtc aatggcgtcc 6420
atccttccat cagatcagaa cgatgaaaaa gtcaagcacg gcatgcatag tatatgtata 6480
gcttgtttta gtgtggcttt gctgagacga atgaaagcaa cggcgggcat atttttcagt 6540
ggctgtagct ttcaggctga aagagacgtg gcatgcaata attcagggaa ttcgtcagcc 6600
aattgaggta gctagtcaac ttgtacattg gtgcgagcaa ttttccgcac tcaggagggc 6660
tagtttgaga gtccaaaaac tataggagat taaagaggct aaaatcctct ccttatttaa 6720
ttttaaataa gtagtgtatt tgtattttaa ctcctccaac ccttccgatt ttatggctct 6780
caaactagca ttcagtctaa tgcatgcatg cttggctaga ggtcgtatgg ggttgttaat 6840
agcatagcta gctacaagtt aaccgggtct tttatattta ataaggacag gcaaagtatt 6900
acttacaaat aaagaataaa gctaggacga actcgtggat tattactaaa tcgaaatgga 6960
cgtaatattc caggcaagaa taattgttcg atcaggagac aagtggggca ttggaccggt 7020
tcttgcaagc aagagcctat ggcgtggtga cacggcgcgt tgcccataca tcatgcctcc 7080
atcgatgatc catcctcact tgctataaaa agaggtgtcc atggtgctca agctcagcca 7140
agcaaataag acgacttgtt tcattgattc ttcaagagat cgagcttctt ttgcaccaca 7200
aggtcgag 7208

<210> 6
<211> 1801
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (1)..(1791)
<223> Maize optimized modified cry3A054 coding sequence.

<220>
<221> misc_feature
<222> (322)..(333)
<223> cathepsin G recognition site coding sequence

<400> 6

```
atg acg gcc gac aac aac acc gag gcc ctg gac agc agc acc acc aag      48
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1               5                  10                  15

gac gtg atc cag aag ggc atc agc gtg gtg ggc gac ctg ctg ggc gtg      96
Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
            20                  25                  30

gtg ggc ttc ccc ttc ggc ggc gcc ctg gtg agc ttc tac acc aac ttc     144
Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
        35                  40                  45

ctg aac acc atc tgg ccc agc gag gac ccc tgg aag gcc ttc atg gag     192
Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
    50                  55                  60

cag gtg gag gcc ctg atg gac cag aag atc gcc gac tac gcc aag aac     240
Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65                  70                  75                  80

aag gca ctg gcc gag cta cag ggc ctc cag aac aac gtg gag gac tat     288
Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
                85                  90                  95

gtg agc gcc ctg agc agc tgg cag aag aac ccc gct gca ccg ttc ccc     336
Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Pro
            100                 105                 110

cac agc cag ggc cgc atc cgc gag ctg ttc agc cag gcc gag agc cac     384
His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser His
        115                 120                 125

ttc cgc aac agc atg ccc agc ttc gcc atc agc ggc tac gag gtg ctg     432
Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val Leu
    130                 135                 140

ttc ctg acc acc tac gcc cag gcc gcc aac acc cac ctg ttc ctg ctg     480
Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu Leu
145                 150                 155                 160

aag gac gcc caa atc tac gga gag gag tgg ggc tac gag aag gag gac     528
Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu Asp
                165                 170                 175

atc gcc gag ttc tac aag cgc cag ctg aag ctg acc cag gag tac acc     576
Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr Thr
            180                 185                 190

gac cac tgc gtg aag tgg tac aac gtg ggt cta gac aag ctc cgc ggc     624
Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg Gly
        195                 200                 205
```

```
agc agc tac gag agc tgg gtg aac ttc aac cgc tac cgc cgc gag atg      672
Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu Met
    210                 215                 220

acc ctg acc gtg ctg gac ctg atc gcc ctg ttc ccc ctg tac gac gtg      720
Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp Val
225                 230                 235                 240

cgc ctg tac ccc aag gag gtg aag acc gag ctg acc cgc gac gtg ctg      768
Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val Leu
                245                 250                 255

acc gac ccc atc gtg ggc gtg aac aac ctg cgc ggc tac ggc acc acc      816
Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr Thr
            260                 265                 270

ttc agc aac atc gag aac tac atc cgc aag ccc cac ctg ttc gac tac      864
Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp Tyr
        275                 280                 285

ctg cac cgc atc cag ttc cac acg cgt ttc cag ccc ggc tac tac ggc      912
Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr Gly
    290                 295                 300

aac gac agc ttc aac tac tgg agc ggc aac tac gtg agc acc cgc ccc      960
Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg Pro
305                 310                 315                 320

agc atc ggc agc aac gac atc atc acc agc ccc ttc tac ggc aac aag     1008
Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn Lys
                325                 330                 335

agc agc gag ccc gtg cag aac ctt gag ttc aac ggc gag aag gtg tac     1056
Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val Tyr
            340                 345                 350

cgc gcc gtg gct aac acc aac ctg gcc gtg tgg ccc tct gca gtg tac     1104
Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val Tyr
        355                 360                 365

agc ggc gtg acc aag gtg gag ttc agc cag tac aac gac cag acc gac     1152
Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr Asp
    370                 375                 380

gag gcc agc acc cag acc tac gac agc aag cgc aac gtg ggc gcc gtg     1200
Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala Val
385                 390                 395                 400

agc tgg gac agc atc gac cag ctg ccc ccc gag acc acc gac gag ccc     1248
Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu Pro
                405                 410                 415

ctg gag aag ggc tac agc cac cag ctg aac tac gtg atg tgc ttc ctg     1296
Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe Leu
```

```
            420                 425                 430

atg cag ggc agc cgc ggc acc atc ccc gtg ctg acc tgg acc cac aag     1344
Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His Lys
        435                 440                 445

agc gtc gac ttc ttc aac atg atc gac agc aag aag atc acc cag ctg     1392
Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln Leu
    450                 455                 460

ccc ctg gtg aag gcc tac aag ctc cag agc ggc gcc agc gtg gtg gca     1440
Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val Ala
465                 470                 475                 480

ggc ccc cgc ttc acc ggc ggc gac atc atc cag tgc acc gag aac ggc     1488
Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn Gly
                485                 490                 495

agc gcc gcc acc atc tac gtg acc ccc gac gtg agc tac agc cag aag     1536
Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln Lys
            500                 505                 510

tac cgc gcc cgc atc cac tac gcc agc acc agc cag atc acc ttc acc     1584
Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe Thr
        515                 520                 525

ctg agc ctg gac ggg gcc ccc ttc aac caa tac tac ttc gac aag acc     1632
Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp Lys Thr
    530                 535                 540

atc aac aag ggc gac acc ctg acc tac aac agc ttc aac ctg gcc agc     1680
Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala Ser
545                 550                 555                 560

ttc agc acc cct ttc gag ctg agc ggc aac aac ctc cag atc ggc gtg     1728
Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly Val
                565                 570                 575

acc ggc ctg agc gcc ggc gac aag gtg tac atc gac aag atc gag ttc     1776
Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu Phe
            580                 585                 590

atc ccc gtg aac tag atctgagctc                                     1801
Ile Pro Val Asn
        595
```

<210> 7
<211> 596
<212> PRT
<213> Artificial Sequence

<220>
<221> misc_feature

<222> (322)..(333)
<223> cathepsin G recognition site coding sequence

<400> 7

```
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1               5                  10                  15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
            20                  25                  30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
        35                  40                  45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
    50                  55                  60

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

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
                85                  90                  95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Pro
            100                 105                 110

His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser His
        115                 120                 125

Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val Leu
    130                 135                 140

Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu Leu
145                 150                 155                 160

Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu Asp
                165                 170                 175

Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr Thr
            180                 185                 190

Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg Gly
        195                 200                 205

Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu Met
    210                 215                 220

Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp Val
225                 230                 235                 240

Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val Leu
                245                 250                 255

Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr Thr
            260                 265                 270
```

Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp Tyr
275 280 285

Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr Gly
290 295 300

Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg Pro
305 310 315 320

Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn Lys
325 330 335

Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val Tyr
340 345 350

Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val Tyr
355 360 365

Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr Asp
370 375 380

Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala Val
385 390 395 400

Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu Pro
405 410 415

Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe Leu
420 425 430

Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His Lys
435 440 445

Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln Leu
450 455 460

Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val Ala
465 470 475 480

Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn Gly
485 490 495

Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln Lys
500 505 510

Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe Thr
515 520 525

Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp Lys Thr
530 535 540

Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala Ser
545 550 555 560

```
Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly Val
                565                 570                 575

Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu Phe
            580                 585                 590

Ile Pro Val Asn
        595


<210>  8
<211>  1807
<212>  DNA
<213>  Artificial Sequence

<220>
<221>  CDS
<222>  (1)..(1806)
<223>  Maize optimized  modified cry3A055 coding sequence.

<220>
<221>  misc_feature
<222>  (322)..(333)
<223>  Cthepsin G recognition site coding sequence.

<400>  8
atg acg gcc gac aac aac acc gag gcc ctg gac agc agc acc acc aag      48
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1               5                   10                  15

gac gtg atc cag aag ggc atc agc gtg gtg ggc gac ctg ctg ggc gtg      96
Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
            20                  25                  30

gtg ggc ttc ccc ttc ggc ggc gcc ctg gtg agc ttc tac acc aac ttc     144
Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
        35                  40                  45

ctg aac acc atc tgg ccc agc gag gac ccc tgg aag gcc ttc atg gag     192
Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
    50                  55                  60

cag gtg gag gcc ctg atg gac cag aag atc gcc gac tac gcc aag aac     240
Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65                  70                  75                  80

aag gca ctg gcc gag cta cag ggc ctc cag aac aac gtg gag gac tat     288
Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
                85                  90                  95

gtg agc gcc ctg agc agc tgg cag aag aac ccc gct gca ccg ttc cgc     336
Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Arg
            100                 105                 110
```

```
aac ccc cac agc cag ggc cgc atc cgc gag ctg ttc agc cag gcc gag      384
Asn Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu
        115                 120                 125

agc cac ttc cgc aac agc atg ccc agc ttc gcc atc agc ggc tac gag      432
Ser His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu
    130                 135                 140

gtg ctg ttc ctg acc acc tac gcc cag gcc gcc aac acc cac ctg ttc      480
Val Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe
145                 150                 155                 160

ctg ctg aag gac gcc caa atc tac gga gag gag tgg ggc tac gag aag      528
Leu Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys
                165                 170                 175

gag gac atc gcc gag ttc tac aag cgc cag ctg aag ctg acc cag gag      576
Glu Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu
            180                 185                 190

tac acc gac cac tgc gtg aag tgg tac aac gtg ggt cta gac aag ctc      624
Tyr Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu
        195                 200                 205

cgc ggc agc agc tac gag agc tgg gtg aac ttc aac cgc tac cgc cgc      672
Arg Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg
    210                 215                 220

gag atg acc ctg acc gtg ctg gac ctg atc gcc ctg ttc ccc ctg tac      720
Glu Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr
225                 230                 235                 240

gac gtg cgc ctg tac ccc aag gag gtg aag acc gag ctg acc cgc gac      768
Asp Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp
                245                 250                 255

gtg ctg acc gac ccc atc gtg ggc gtg aac aac ctg cgc ggc tac ggc      816
Val Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly
            260                 265                 270

acc acc ttc agc aac atc gag aac tac atc cgc aag ccc cac ctg ttc      864
Thr Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe
        275                 280                 285

gac tac ctg cac cgc atc cag ttc cac acg cgt ttc cag ccc ggc tac      912
Asp Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr
    290                 295                 300

tac ggc aac gac agc ttc aac tac tgg agc ggc aac tac gtg agc acc      960
Tyr Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr
305                 310                 315                 320

cgc ccc agc atc ggc agc aac gac atc atc acc agc ccc ttc tac ggc     1008
```

```
Arg Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly
                325                 330                 335

aac aag agc agc gag ccc gtg cag aac ctt gag ttc aac ggc gag aag   1056
Asn Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys
            340                 345                 350

gtg tac cgc gcc gtg gct aac acc aac ctg gcc gtg tgg ccc tct gca   1104
Val Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala
        355                 360                 365

gtg tac agc ggc gtg acc aag gtg gag ttc agc cag tac aac gac cag   1152
Val Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln
    370                 375                 380

acc gac gag gcc agc acc cag acc tac gac agc aag cgc aac gtg ggc   1200
Thr Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly
385                 390                 395                 400

gcc gtg agc tgg gac agc atc gac cag ctg ccc ccc gag acc acc gac   1248
Ala Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp
                405                 410                 415

gag ccc ctg gag aag ggc tac agc cac cag ctg aac tac gtg atg tgc   1296
Glu Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys
            420                 425                 430

ttc ctg atg cag ggc agc cgc ggc acc atc ccc gtg ctg acc tgg acc   1344
Phe Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr
        435                 440                 445

cac aag agc gtc gac ttc ttc aac atg atc gac agc aag aag atc acc   1392
His Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr
    450                 455                 460

cag ctg ccc ctg gtg aag gcc tac aag ctc cag agc ggc gcc agc gtg   1440
Gln Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val
465                 470                 475                 480

gtg gca ggc ccc cgc ttc acc ggc ggc gac atc atc cag tgc acc gag   1488
Val Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu
                485                 490                 495

aac ggc agc gcc gcc acc atc tac gtg acc ccc gac gtg agc tac agc   1536
Asn Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser
            500                 505                 510

cag aag tac cgc gcc cgc atc cac tac gcc agc acc agc cag atc acc   1584
Gln Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr
        515                 520                 525

ttc acc ctg agc ctg gac ggg gcc ccc ttc aac caa tac tac ttc gac   1632
Phe Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp
    530                 535                 540
```

```
aag acc atc aac aag ggc gac acc ctg acc tac aac agc ttc aac ctg     1680
Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu
545                 550                 555                 560

gcc agc ttc agc acc cct ttc gag ctg agc ggc aac aac ctc cag atc     1728
Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile
                565                 570                 575

ggc gtg acc ggc ctg agc gcc ggc gac aag gtg tac atc gac aag atc     1776
Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile
            580                 585                 590

gag ttc atc ccc gtg aac tag atc tga gct c                           1807
Glu Phe Ile Pro Val Asn     Ile     Ala
        595                         600
```

<210> 9
<211> 598
<212> PRT
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (322)..(333)
<223> Cthepsin G recognition site coding sequence.

<400> 9

```
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1               5                   10                  15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
            20                  25                  30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
        35                  40                  45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
    50                  55                  60

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

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
                85                  90                  95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Arg
            100                 105                 110

Asn Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu
        115                 120                 125

Ser His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu
```

```
    130                 135                 140

Val Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe
145                 150                 155                 160

Leu Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys
                165                 170                 175

Glu Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu
            180                 185                 190

Tyr Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu
        195                 200                 205

Arg Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg
    210                 215                 220

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

Asp Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp
                245                 250                 255

Val Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly
            260                 265                 270

Thr Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe
        275                 280                 285

Asp Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr
    290                 295                 300

Tyr Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr
305                 310                 315                 320

Arg Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly
                325                 330                 335

Asn Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys
            340                 345                 350

Val Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala
        355                 360                 365

Val Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln
    370                 375                 380

Thr Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly
385                 390                 395                 400

Ala Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp
                405                 410                 415

Glu Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys
```

420 425 430

Phe Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr
435 440 445

His Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr
450 455 460

Gln Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val
465 470 475 480

Val Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu
485 490 495

Asn Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser
500 505 510

Gln Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr
515 520 525

Phe Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp
530 535 540

Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu
545 550 555 560

Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile
565 570 575

Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile
580 585 590

Glu Phe Ile Pro Val Asn
595

<210> 10
<211> 1818
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (1)..(1818)
<223> Maize optimized modified cry3A085 coding sequence.

<220>
<221> misc_feature
<222> (346)..(357)
<223> Cathepsin G recognition site coding sequence.

<400> 10
atg aac tac aag gag ttc ctc cgc atg acc gcc gac aac aac acc gag 48
Met Asn Tyr Lys Glu Phe Leu Arg Met Thr Ala Asp Asn Asn Thr Glu

```
1               5                   10                  15

gcc ctg gac agc agc acc acc aag gac gtg atc cag aag ggc atc agc      96
Ala Leu Asp Ser Ser Thr Thr Lys Asp Val Ile Gln Lys Gly Ile Ser
            20                  25                  30

gtg gtg ggc gac ctg ctg ggc gtg gtg ggc ttc ccc ttc ggc ggc gcc     144
Val Val Gly Asp Leu Leu Gly Val Val Gly Phe Pro Phe Gly Gly Ala
        35                  40                  45

ctg gtg agc ttc tac acc aac ttc ctg aac acc atc tgg ccc agc gag     192
Leu Val Ser Phe Tyr Thr Asn Phe Leu Asn Thr Ile Trp Pro Ser Glu
    50                  55                  60

gac ccc tgg aag gcc ttc atg gag cag gtg gag gcc ctg atg gac cag     240
Asp Pro Trp Lys Ala Phe Met Glu Gln Val Glu Ala Leu Met Asp Gln
65                  70                  75                  80

aag atc gcc gac tac gcc aag aac aag gca ctg gcc gag cta cag ggc     288
Lys Ile Ala Asp Tyr Ala Lys Asn Lys Ala Leu Ala Glu Leu Gln Gly
                85                  90                  95

ctc cag aac aac gtg gag gac tat gtg agc gcc ctg agc agc tgg cag     336
Leu Gln Asn Asn Val Glu Asp Tyr Val Ser Ala Leu Ser Ser Trp Gln
            100                 105                 110

aag aac ccc gct gca ccg ttc cgc aac ccc cac agc cag ggc cgc atc     384
Lys Asn Pro Ala Ala Pro Phe Arg Asn Pro His Ser Gln Gly Arg Ile
        115                 120                 125

cgc gag ctg ttc agc cag gcc gag agc cac ttc cgc aac agc atg ccc     432
Arg Glu Leu Phe Ser Gln Ala Glu Ser His Phe Arg Asn Ser Met Pro
    130                 135                 140

agc ttc gcc atc agc ggc tac gag gtg ctg ttc ctg acc acc tac gcc     480
Ser Phe Ala Ile Ser Gly Tyr Glu Val Leu Phe Leu Thr Thr Tyr Ala
145                 150                 155                 160

cag gcc gcc aac acc cac ctg ttc ctg ctg aag gac gcc caa atc tac     528
Gln Ala Ala Asn Thr His Leu Phe Leu Leu Lys Asp Ala Gln Ile Tyr
                165                 170                 175

gga gag gag tgg ggc tac gag aag gag gac atc gcc gag ttc tac aag     576
Gly Glu Glu Trp Gly Tyr Glu Lys Glu Asp Ile Ala Glu Phe Tyr Lys
            180                 185                 190

cgc cag ctg aag ctg acc cag gag tac acc gac cac tgc gtg aag tgg     624
Arg Gln Leu Lys Leu Thr Gln Glu Tyr Thr Asp His Cys Val Lys Trp
        195                 200                 205

tac aac gtg ggt cta gac aag ctc cgc ggc agc agc tac gag agc tgg     672
Tyr Asn Val Gly Leu Asp Lys Leu Arg Gly Ser Ser Tyr Glu Ser Trp
    210                 215                 220
```

```
gtg aac ttc aac cgc tac cgc cgc gag atg acc ctg acc gtg ctg gac      720
Val Asn Phe Asn Arg Tyr Arg Arg Glu Met Thr Leu Thr Val Leu Asp
225                 230                 235                 240

ctg atc gcc ctg ttc ccc ctg tac gac gtg cgc ctg tac ccc aag gag      768
Leu Ile Ala Leu Phe Pro Leu Tyr Asp Val Arg Leu Tyr Pro Lys Glu
                245                 250                 255

gtg aag acc gag ctg acc cgc gac gtg ctg acc gac ccc atc gtg ggc      816
Val Lys Thr Glu Leu Thr Arg Asp Val Leu Thr Asp Pro Ile Val Gly
            260                 265                 270

gtg aac aac ctg cgc ggc tac ggc acc acc ttc agc aac atc gag aac      864
Val Asn Asn Leu Arg Gly Tyr Gly Thr Thr Phe Ser Asn Ile Glu Asn
        275                 280                 285

tac atc cgc aag ccc cac ctg ttc gac tac ctg cac cgc atc cag ttc      912
Tyr Ile Arg Lys Pro His Leu Phe Asp Tyr Leu His Arg Ile Gln Phe
    290                 295                 300

cac acg cgt ttc cag ccc ggc tac tac ggc aac gac agc ttc aac tac      960
His Thr Arg Phe Gln Pro Gly Tyr Tyr Gly Asn Asp Ser Phe Asn Tyr
305                 310                 315                 320

tgg agc ggc aac tac gtg agc acc cgc ccc agc atc ggc agc aac gac     1008
Trp Ser Gly Asn Tyr Val Ser Thr Arg Pro Ser Ile Gly Ser Asn Asp
                325                 330                 335

atc atc acc agc ccc ttc tac ggc aac aag agc agc gag ccc gtg cag     1056
Ile Ile Thr Ser Pro Phe Tyr Gly Asn Lys Ser Ser Glu Pro Val Gln
            340                 345                 350

aac ctt gag ttc aac ggc gag aag gtg tac cgc gcc gtg gct aac acc     1104
Asn Leu Glu Phe Asn Gly Glu Lys Val Tyr Arg Ala Val Ala Asn Thr
        355                 360                 365

aac ctg gcc gtg tgg ccc tct gca gtg tac agc ggc gtg acc aag gtg     1152
Asn Leu Ala Val Trp Pro Ser Ala Val Tyr Ser Gly Val Thr Lys Val
    370                 375                 380

gag ttc agc cag tac aac gac cag acc gac gag gcc agc acc cag acc     1200
Glu Phe Ser Gln Tyr Asn Asp Gln Thr Asp Glu Ala Ser Thr Gln Thr
385                 390                 395                 400

tac gac agc aag cgc aac gtg ggc gcc gtg agc tgg gac agc atc gac     1248
Tyr Asp Ser Lys Arg Asn Val Gly Ala Val Ser Trp Asp Ser Ile Asp
                405                 410                 415

cag ctg ccc ccc gag acc acc gac gag ccc ctg gag aag ggc tac agc     1296
Gln Leu Pro Pro Glu Thr Thr Asp Glu Pro Leu Glu Lys Gly Tyr Ser
            420                 425                 430

cac cag ctg aac tac gtg atg tgc ttc ctg atg cag ggc agc cgc ggc     1344
His Gln Leu Asn Tyr Val Met Cys Phe Leu Met Gln Gly Ser Arg Gly
```

```
        435                 440                 445

acc atc ccc gtg ctg acc tgg acc cac aag agc gtc gac ttc ttc aac     1392
Thr Ile Pro Val Leu Thr Trp Thr His Lys Ser Val Asp Phe Phe Asn
    450                 455                 460

atg atc gac agc aag aag atc acc cag ctg ccc ctg gtg aag gcc tac     1440
Met Ile Asp Ser Lys Lys Ile Thr Gln Leu Pro Leu Val Lys Ala Tyr
465                 470                 475                 480

aag ctc cag agc ggc gcc agc gtg gtg gca ggc ccc cgc ttc acc ggc     1488
Lys Leu Gln Ser Gly Ala Ser Val Val Ala Gly Pro Arg Phe Thr Gly
                485                 490                 495

ggc gac atc atc cag tgc acc gag aac ggc agc gcc gcc acc atc tac     1536
Gly Asp Ile Ile Gln Cys Thr Glu Asn Gly Ser Ala Ala Thr Ile Tyr
            500                 505                 510

gtg acc ccc gac gtg agc tac agc cag aag tac cgc gcc cgc atc cac     1584
Val Thr Pro Asp Val Ser Tyr Ser Gln Lys Tyr Arg Ala Arg Ile His
        515                 520                 525

tac gcc agc acc agc cag atc acc ttc acc ctg agc ctg gac ggg gcc     1632
Tyr Ala Ser Thr Ser Gln Ile Thr Phe Thr Leu Ser Leu Asp Gly Ala
    530                 535                 540

ccc ttc aac caa tac tac ttc gac aag acc atc aac aag ggc gac acc     1680
Pro Phe Asn Gln Tyr Tyr Phe Asp Lys Thr Ile Asn Lys Gly Asp Thr
545                 550                 555                 560

ctg acc tac aac agc ttc aac ctg gcc agc ttc agc acc cct ttc gag     1728
Leu Thr Tyr Asn Ser Phe Asn Leu Ala Ser Phe Ser Thr Pro Phe Glu
                565                 570                 575

ctg agc ggc aac aac ctc cag atc ggc gtg acc ggc ctg agc gcc ggc     1776
Leu Ser Gly Asn Asn Leu Gln Ile Gly Val Thr Gly Leu Ser Ala Gly
            580                 585                 590

gac aag gtg tac atc gac aag atc gag ttc atc ccc gtg aac             1818
Asp Lys Val Tyr Ile Asp Lys Ile Glu Phe Ile Pro Val Asn
        595                 600                 605
```

<210> 11
<211> 606
<212> PRT
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (346)..(357)
<223> Cathepsin G recognition site coding sequence.

<400> 11

```
Met Asn Tyr Lys Glu Phe Leu Arg Met Thr Ala Asp Asn Asn Thr Glu
1               5                  10                  15

Ala Leu Asp Ser Ser Thr Thr Lys Asp Val Ile Gln Lys Gly Ile Ser
            20                  25                  30

Val Val Gly Asp Leu Leu Gly Val Val Gly Phe Pro Phe Gly Gly Ala
        35                  40                  45

Leu Val Ser Phe Tyr Thr Asn Phe Leu Asn Thr Ile Trp Pro Ser Glu
    50                  55                  60

Asp Pro Trp Lys Ala Phe Met Glu Gln Val Glu Ala Leu Met Asp Gln
65                  70                  75                  80

Lys Ile Ala Asp Tyr Ala Lys Asn Lys Ala Leu Ala Glu Leu Gln Gly
                85                  90                  95

Leu Gln Asn Asn Val Glu Asp Tyr Val Ser Ala Leu Ser Ser Trp Gln
            100                 105                 110

Lys Asn Pro Ala Ala Pro Phe Arg Asn Pro His Ser Gln Gly Arg Ile
        115                 120                 125

Arg Glu Leu Phe Ser Gln Ala Glu Ser His Phe Arg Asn Ser Met Pro
    130                 135                 140

Ser Phe Ala Ile Ser Gly Tyr Glu Val Leu Phe Leu Thr Thr Tyr Ala
145                 150                 155                 160

Gln Ala Ala Asn Thr His Leu Phe Leu Leu Lys Asp Ala Gln Ile Tyr
                165                 170                 175

Gly Glu Glu Trp Gly Tyr Glu Lys Glu Asp Ile Ala Glu Phe Tyr Lys
            180                 185                 190

Arg Gln Leu Lys Leu Thr Gln Glu Tyr Thr Asp His Cys Val Lys Trp
        195                 200                 205

Tyr Asn Val Gly Leu Asp Lys Leu Arg Gly Ser Ser Tyr Glu Ser Trp
    210                 215                 220

Val Asn Phe Asn Arg Tyr Arg Arg Glu Met Thr Leu Thr Val Leu Asp
225                 230                 235                 240

Leu Ile Ala Leu Phe Pro Leu Tyr Asp Val Arg Leu Tyr Pro Lys Glu
                245                 250                 255

Val Lys Thr Glu Leu Thr Arg Asp Val Leu Thr Asp Pro Ile Val Gly
            260                 265                 270

Val Asn Asn Leu Arg Gly Tyr Gly Thr Thr Phe Ser Asn Ile Glu Asn
        275                 280                 285
```

```
Tyr Ile Arg Lys Pro His Leu Phe Asp Tyr Leu His Arg Ile Gln Phe
    290                 295                 300

His Thr Arg Phe Gln Pro Gly Tyr Tyr Gly Asn Asp Ser Phe Asn Tyr
305                 310                 315                 320

Trp Ser Gly Asn Tyr Val Ser Thr Arg Pro Ser Ile Gly Ser Asn Asp
                325                 330                 335

Ile Ile Thr Ser Pro Phe Tyr Gly Asn Lys Ser Ser Glu Pro Val Gln
            340                 345                 350

Asn Leu Glu Phe Asn Gly Glu Lys Val Tyr Arg Ala Val Ala Asn Thr
        355                 360                 365

Asn Leu Ala Val Trp Pro Ser Ala Val Tyr Ser Gly Val Thr Lys Val
    370                 375                 380

Glu Phe Ser Gln Tyr Asn Asp Gln Thr Asp Glu Ala Ser Thr Gln Thr
385                 390                 395                 400

Tyr Asp Ser Lys Arg Asn Val Gly Ala Val Ser Trp Asp Ser Ile Asp
                405                 410                 415

Gln Leu Pro Pro Glu Thr Thr Asp Glu Pro Leu Glu Lys Gly Tyr Ser
            420                 425                 430

His Gln Leu Asn Tyr Val Met Cys Phe Leu Met Gln Gly Ser Arg Gly
        435                 440                 445

Thr Ile Pro Val Leu Thr Trp Thr His Lys Ser Val Asp Phe Phe Asn
    450                 455                 460

Met Ile Asp Ser Lys Lys Ile Thr Gln Leu Pro Leu Val Lys Ala Tyr
465                 470                 475                 480

Lys Leu Gln Ser Gly Ala Ser Val Val Ala Gly Pro Arg Phe Thr Gly
                485                 490                 495

Gly Asp Ile Ile Gln Cys Thr Glu Asn Gly Ser Ala Ala Thr Ile Tyr
            500                 505                 510

Val Thr Pro Asp Val Ser Tyr Ser Gln Lys Tyr Arg Ala Arg Ile His
        515                 520                 525

Tyr Ala Ser Thr Ser Gln Ile Thr Phe Thr Leu Ser Leu Asp Gly Ala
    530                 535                 540

Pro Phe Asn Gln Tyr Tyr Phe Asp Lys Thr Ile Asn Lys Gly Asp Thr
545                 550                 555                 560

Leu Thr Tyr Asn Ser Phe Asn Leu Ala Ser Phe Ser Thr Pro Phe Glu
                565                 570                 575
```

Leu Ser Gly Asn Asn Leu Gln Ile Gly Val Thr Gly Leu Ser Ala Gly
580 585 590

Asp Lys Val Tyr Ile Asp Lys Ile Glu Phe Ile Pro Val Asn
595 600 605

<210> 12
<211> 1794
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (1)..(1794)
<223> Maize optimized modified cry3A082 coding sequence.

<220>
<221> misc_feature
<222> (1609)..(1620)
<223> Cathepsin G recognition site coding sequence

<400> 12
atg acg gcc gac aac aac acc gag gcc ctg gac agc agc acc acc aag 48
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

gac gtg atc cag aag ggc atc agc gtg gtg ggc gac ctg ctg ggc gtg 96
Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

gtg ggc ttc ccc ttc ggc ggc gcc ctg gtg agc ttc tac acc aac ttc 144
Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

ctg aac acc atc tgg ccc agc gag gac ccc tgg aag gcc ttc atg gag 192
Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

cag gtg gag gcc ctg atg gac cag aag atc gcc gac tac gcc aag aac 240
Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65 70 75 80

aag gca ctg gcc gag cta cag ggc ctc cag aac aac gtg gag gac tat 288
Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

gtg agc gcc ctg agc agc tgg cag aag aac ccc gtc tcg agc cgc aac 336
Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn
100 105 110

ccc cac agc cag ggc cgc atc cgc gag ctg ttc agc cag gcc gag agc 384
Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser
115 120 125

```
cac ttc cgc aac agc atg ccc agc ttc gcc atc agc ggc tac gag gtg      432
His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val
    130                 135                 140

ctg ttc ctg acc acc tac gcc cag gcc gcc aac acc cac ctg ttc ctg      480
Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu
145                 150                 155                 160

ctg aag gac gcc caa atc tac gga gag gag tgg ggc tac gag aag gag      528
Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu
                165                 170                 175

gac atc gcc gag ttc tac aag cgc cag ctg aag ctg acc cag gag tac      576
Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr
            180                 185                 190

acc gac cac tgc gtg aag tgg tac aac gtg ggt cta gac aag ctc cgc      624
Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg
        195                 200                 205

ggc agc agc tac gag agc tgg gtg aac ttc aac cgc tac cgc cgc gag      672
Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu
    210                 215                 220

atg acc ctg acc gtg ctg gac ctg atc gcc ctg ttc ccc ctg tac gac      720
Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp
225                 230                 235                 240

gtg cgc ctg tac ccc aag gag gtg aag acc gag ctg acc cgc gac gtg      768
Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val
                245                 250                 255

ctg acc gac ccc atc gtg ggc gtg aac aac ctg cgc ggc tac ggc acc      816
Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr
            260                 265                 270

acc ttc agc aac atc gag aac tac atc cgc aag ccc cac ctg ttc gac      864
Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp
        275                 280                 285

tac ctg cac cgc atc cag ttc cac acg cgt ttc cag ccc ggc tac tac      912
Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr
    290                 295                 300

ggc aac gac agc ttc aac tac tgg agc ggc aac tac gtg agc acc cgc      960
Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg
305                 310                 315                 320

ccc agc atc ggc agc aac gac atc atc acc agc ccc ttc tac ggc aac     1008
Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn
                325                 330                 335

aag agc agc gag ccc gtg cag aac ctt gag ttc aac ggc gag aag gtg     1056
```

```
Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val
            340                 345                 350

tac cgc gcc gtg gct aac acc aac ctg gcc gtg tgg ccc tct gca gtg    1104
Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val
        355                 360                 365

tac agc ggc gtg acc aag gtg gag ttc agc cag tac aac gac cag acc    1152
Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr
    370                 375                 380

gac gag gcc agc acc cag acc tac gac agc aag cgc aac gtg ggc gcc    1200
Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala
385                 390                 395                 400

gtg agc tgg gac agc atc gac cag ctg ccc ccc gag acc acc gac gag    1248
Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu
                405                 410                 415

ccc ctg gag aag ggc tac agc cac cag ctg aac tac gtg atg tgc ttc    1296
Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe
            420                 425                 430

ctg atg cag ggc agc cgc ggc acc atc ccc gtg ctg acc tgg acc cac    1344
Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His
        435                 440                 445

aag agc gtc gac ttc ttc aac atg atc gac agc aag aag atc acc cag    1392
Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln
    450                 455                 460

ctg ccc ctg gtg aag gcc tac aag ctc cag agc ggc gcc agc gtg gtg    1440
Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val
465                 470                 475                 480

gca ggc ccc cgc ttc acc ggc ggc gac atc atc cag tgc acc gag aac    1488
Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn
                485                 490                 495

ggc agc gcc gcc acc atc tac gtg acc ccc gac gtg agc tac agc cag    1536
Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln
            500                 505                 510

aag tac cgc gcc cgc atc cac tac gcc agc acc agc cag atc acc ttc    1584
Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe
        515                 520                 525

acc ctg agc ctg gac ggg gcc ccc gct gca ccg ttc tac ttc gac aag    1632
Thr Leu Ser Leu Asp Gly Ala Pro Ala Ala Pro Phe Tyr Phe Asp Lys
    530                 535                 540

acc atc aac aag ggc gac acc ctg acc tac aac agc ttc aac ctg gcc    1680
Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala
545                 550                 555                 560
```

```
agc ttc agc acc cct ttc gag ctg agc ggc aac aac ctc cag atc ggc      1728
Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly
                565                 570                 575

gtg acc ggc ctg agc gcc ggc gac aag gtg tac atc gac aag atc gag      1776
Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu
            580                 585                 590

ttc atc ccc gtg aac tag                                              1794
Phe Ile Pro Val Asn
        595
```

<210> 13
<211> 597
<212> PRT
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1609)..(1620)
<223> Cathepsin G recognition site coding sequence

<400> 13

```
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1               5                   10                  15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
            20                  25                  30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
        35                  40                  45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
    50                  55                  60

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

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
                85                  90                  95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn
            100                 105                 110

Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser
        115                 120                 125

His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val
    130                 135                 140

Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu
145                 150                 155                 160
```

Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu
165 170 175

Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr
180 185 190

Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg
195 200 205

Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu
210 215 220

Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp
225 230 235 240

Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val
245 250 255

Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr
260 265 270

Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp
275 280 285

Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr
290 295 300

Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg
305 310 315 320

Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn
325 330 335

Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val
340 345 350

Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val
355 360 365

Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr
370 375 380

Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala
385 390 395 400

Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu
405 410 415

Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe
420 425 430

Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His
435 440 445

```
Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln
    450                 455                 460

Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val
465                 470                 475                 480

Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn
                485                 490                 495

Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln
            500                 505                 510

Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe
        515                 520                 525

Thr Leu Ser Leu Asp Gly Ala Pro Ala Ala Pro Phe Tyr Phe Asp Lys
    530                 535                 540

Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala
545                 550                 555                 560

Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly
                565                 570                 575

Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu
            580                 585                 590

Phe Ile Pro Val Asn
        595
```

<210> 14
<211> 1816
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (1)..(1812)
<223> Maize optimized modified cry3A058 coding sequence.

<220>
<221> misc_feature
<222> (1621)..(1632)
<223> Cathepsin G recognition site coding sequence

<400> 14

```
atg acg gcc gac aac aac acc gag gcc ctg gac agc agc acc acc aag      48
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1               5                   10                  15

gac gtg atc cag aag ggc atc agc gtg gtg ggc gac ctg ctg ggc gtg      96
Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
```

```
                20                  25                  30

gtg ggc ttc ccc ttc ggc ggc gcc ctg gtg agc ttc tac acc aac ttc      144
Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
        35                  40                  45

ctg aac acc atc tgg ccc agc gag gac ccc tgg aag gcc ttc atg gag      192
Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
    50                  55                  60

cag gtg gag gcc ctg atg gac cag aag atc gcc gac tac gcc aag aac      240
Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65                  70                  75                  80

aag gca ctg gcc gag cta cag ggc ctc cag aac aac gtg gag gac tat      288
Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
                85                  90                  95

gtg agc gcc ctg agc agc tgg cag aag aac ccc gtc tcg agc cgc aac      336
Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn
            100                 105                 110

ccc cac agc cag ggc cgc atc cgc gag ctg ttc agc cag gcc gag agc      384
Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser
        115                 120                 125

cac ttc cgc aac agc atg ccc agc ttc gcc atc agc ggc tac gag gtg      432
His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val
    130                 135                 140

ctg ttc ctg acc acc tac gcc cag gcc gcc aac acc cac ctg ttc ctg      480
Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu
145                 150                 155                 160

ctg aag gac gcc caa atc tac gga gag gag tgg ggc tac gag aag gag      528
Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu
                165                 170                 175

gac atc gcc gag ttc tac aag cgc cag ctg aag ctg acc cag gag tac      576
Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr
            180                 185                 190

acc gac cac tgc gtg aag tgg tac aac gtg ggt cta gac aag ctc cgc      624
Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg
        195                 200                 205

ggc agc agc tac gag agc tgg gtg aac ttc aac cgc tac cgc cgc gag      672
Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu
    210                 215                 220

atg acc ctg acc gtg ctg gac ctg atc gcc ctg ttc ccc ctg tac gac      720
Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp
225                 230                 235                 240
```

```
gtg cgc ctg tac ccc aag gag gtg aag acc gag ctg acc cgc gac gtg      768
Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val
                245                 250                 255

ctg acc gac ccc atc gtg ggc gtg aac aac ctg cgc ggc tac ggc acc      816
Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr
            260                 265                 270

acc ttc agc aac atc gag aac tac atc cgc aag ccc cac ctg ttc gac      864
Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp
        275                 280                 285

tac ctg cac cgc atc cag ttc cac acg cgt ttc cag ccc ggc tac tac      912
Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr
    290                 295                 300

ggc aac gac agc ttc aac tac tgg agc ggc aac tac gtg agc acc cgc      960
Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg
305                 310                 315                 320

ccc agc atc ggc agc aac gac atc atc acc agc ccc ttc tac ggc aac     1008
Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn
                325                 330                 335

aag agc agc gag ccc gtg cag aac ctt gag ttc aac ggc gag aag gtg     1056
Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val
            340                 345                 350

tac cgc gcc gtg gct aac acc aac ctg gcc gtg tgg ccc tct gca gtg     1104
Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val
        355                 360                 365

tac agc ggc gtg acc aag gtg gag ttc agc cag tac aac gac cag acc     1152
Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr
    370                 375                 380

gac gag gcc agc acc cag acc tac gac agc aag cgc aac gtg ggc gcc     1200
Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala
385                 390                 395                 400

gtg agc tgg gac agc atc gac cag ctg ccc ccc gag acc acc gac gag     1248
Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu
                405                 410                 415

ccc ctg gag aag ggc tac agc cac cag ctg aac tac gtg atg tgc ttc     1296
Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe
            420                 425                 430

ctg atg cag ggc agc cgc ggc acc atc ccc gtg ctg acc tgg acc cac     1344
Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His
        435                 440                 445

aag agc gtc gac ttc ttc aac atg atc gac agc aag aag atc acc cag     1392
Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln
```

```
    450                 455                 460

ctg ccc ctg gtg aag gcc tac aag ctc cag agc ggc gcc agc gtg gtg     1440
Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val
465                 470                 475                 480

gca ggc ccc cgc ttc acc ggc ggc gac atc atc cag tgc acc gag aac     1488
Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn
                485                 490                 495

ggc agc gcc gcc acc atc tac gtg acc ccc gac gtg agc tac agc cag     1536
Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln
            500                 505                 510

aag tac cgc gcc cgc atc cac tac gcc agc acc agc cag atc acc ttc     1584
Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe
        515                 520                 525

acc ctg agc ctg gac ggg gcc ccc ttc aac caa tac gct gca ccg ttc     1632
Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Ala Ala Pro Phe
    530                 535                 540

tac ttc gac aag acc atc aac aag ggc gac acc ctg acc tac aac agc     1680
Tyr Phe Asp Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser
545                 550                 555                 560

ttc aac ctg gcc agc ttc agc acc cct ttc gag ctg agc ggc aac aac     1728
Phe Asn Leu Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn
                565                 570                 575

ctc cag atc ggc gtg acc ggc ctg agc gcc ggc gac aag gtg tac atc     1776
Leu Gln Ile Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile
            580                 585                 590

gac aag atc gag ttc atc ccc gtg aac tag atc tga gctc                1816
Asp Lys Ile Glu Phe Ile Pro Val Asn      Ile
        595                 600
```

<210> 15
<211> 601
<212> PRT
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1621)..(1632)
<223> Cathepsin G recognition site coding sequence

<400> 15

```
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1               5                   10                  15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
```

```
            20                  25                  30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
        35                  40                  45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
    50                  55                  60

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

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
                85                  90                  95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn
            100                 105                 110

Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser
        115                 120                 125

His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val
    130                 135                 140

Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu
145                 150                 155                 160

Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu
                165                 170                 175

Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr
            180                 185                 190

Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg
        195                 200                 205

Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu
    210                 215                 220

Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp
225                 230                 235                 240

Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val
                245                 250                 255

Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr
            260                 265                 270

Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp
        275                 280                 285

Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr
    290                 295                 300

Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg
```

305 310 315 320

Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn
325 330 335

Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val
340 345 350

Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val
355 360 365

Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr
370 375 380

Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala
385 390 395 400

Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu
405 410 415

Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe
420 425 430

Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His
435 440 445

Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln
450 455 460

Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val
465 470 475 480

Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn
485 490 495

Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln
500 505 510

Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe
515 520 525

Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Ala Ala Pro Phe
530 535 540

Tyr Phe Asp Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser
545 550 555 560

Phe Asn Leu Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn
565 570 575

Leu Gln Ile Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile
580 585 590

Asp Lys Ile Glu Phe Ile Pro Val Asn

595 600

<210> 16
<211> 1813
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (1)..(1812)
<223> Maize optimized modified cry3A057 coding sequence.

<220>
<221> misc_feature
<222> (322)..(333)
<223> Cathepsin G recognition site coding sequence

<220>
<221> misc_feature
<222> (1618)..(1629)
<223> Cathepsin G recognition site coding sequence

<400> 16

```
atg acg gcc gac aac aac acc gag gcc ctg gac agc agc acc acc aag      48
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1               5                  10                  15

gac gtg atc cag aag ggc atc agc gtg gtg ggc gac ctg ctg ggc gtg      96
Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
            20                  25                  30

gtg ggc ttc ccc ttc ggc ggc gcc ctg gtg agc ttc tac acc aac ttc     144
Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
        35                  40                  45

ctg aac acc atc tgg ccc agc gag gac ccc tgg aag gcc ttc atg gag     192
Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
    50                  55                  60

cag gtg gag gcc ctg atg gac cag aag atc gcc gac tac gcc aag aac     240
Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65                  70                  75                  80

aag gca ctg gcc gag cta cag ggc ctc cag aac aac gtg gag gac tat     288
Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
                85                  90                  95

gtg agc gcc ctg agc agc tgg cag aag aac ccc gct gca ccg ttc ccc     336
Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Pro
            100                 105                 110

cac agc cag ggc cgc atc cgc gag ctg ttc agc cag gcc gag agc cac     384
His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser His
        115                 120                 125
```

```
ttc cgc aac agc atg ccc agc ttc gcc atc agc ggc tac gag gtg ctg      432
Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val Leu
    130                 135                 140

ttc ctg acc acc tac gcc cag gcc gcc aac acc cac ctg ttc ctg ctg      480
Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu Leu
145                 150                 155                 160

aag gac gcc caa atc tac gga gag gag tgg ggc tac gag aag gag gac      528
Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu Asp
                165                 170                 175

atc gcc gag ttc tac aag cgc cag ctg aag ctg acc cag gag tac acc      576
Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr Thr
            180                 185                 190

gac cac tgc gtg aag tgg tac aac gtg ggt cta gac aag ctc cgc ggc      624
Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg Gly
        195                 200                 205

agc agc tac gag agc tgg gtg aac ttc aac cgc tac cgc cgc gag atg      672
Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu Met
    210                 215                 220

acc ctg acc gtg ctg gac ctg atc gcc ctg ttc ccc ctg tac gac gtg      720
Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp Val
225                 230                 235                 240

cgc ctg tac ccc aag gag gtg aag acc gag ctg acc cgc gac gtg ctg      768
Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val Leu
                245                 250                 255

acc gac ccc atc gtg ggc gtg aac aac ctg cgc ggc tac ggc acc acc      816
Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr Thr
            260                 265                 270

ttc agc aac atc gag aac tac atc cgc aag ccc cac ctg ttc gac tac      864
Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp Tyr
        275                 280                 285

ctg cac cgc atc cag ttc cac acg cgt ttc cag ccc ggc tac tac ggc      912
Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr Gly
    290                 295                 300

aac gac agc ttc aac tac tgg agc ggc aac tac gtg agc acc cgc ccc      960
Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg Pro
305                 310                 315                 320

agc atc ggc agc aac gac atc atc acc agc ccc ttc tac ggc aac aag     1008
Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn Lys
                325                 330                 335

agc agc gag ccc gtg cag aac ctt gag ttc aac ggc gag aag gtg tac     1056
```

```
Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val Tyr
            340                 345                 350

cgc gcc gtg gct aac acc aac ctg gcc gtg tgg ccc tct gca gtg tac    1104
Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val Tyr
        355                 360                 365

agc ggc gtg acc aag gtg gag ttc agc cag tac aac gac cag acc gac    1152
Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr Asp
    370                 375                 380

gag gcc agc acc cag acc tac gac agc aag cgc aac gtg ggc gcc gtg    1200
Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala Val
385                 390                 395                 400

agc tgg gac agc atc gac cag ctg ccc ccc gag acc acc gac gag ccc    1248
Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu Pro
                405                 410                 415

ctg gag aag ggc tac agc cac cag ctg aac tac gtg atg tgc ttc ctg    1296
Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe Leu
            420                 425                 430

atg cag ggc agc cgc ggc acc atc ccc gtg ctg acc tgg acc cac aag    1344
Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His Lys
        435                 440                 445

agc gtc gac ttc ttc aac atg atc gac agc aag aag atc acc cag ctg    1392
Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln Leu
    450                 455                 460

ccc ctg gtg aag gcc tac aag ctc cag agc ggc gcc agc gtg gtg gca    1440
Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val Ala
465                 470                 475                 480

ggc ccc cgc ttc acc ggc ggc gac atc atc cag tgc acc gag aac ggc    1488
Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn Gly
                485                 490                 495

agc gcc gcc acc atc tac gtg acc ccc gac gtg agc tac agc cag aag    1536
Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln Lys
            500                 505                 510

tac cgc gcc cgc atc cac tac gcc agc acc agc cag atc acc ttc acc    1584
Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe Thr
        515                 520                 525

ctg agc ctg gac ggg gcc ccc ttc aac caa tac gct gca ccg ttc tac    1632
Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Ala Ala Pro Phe Tyr
    530                 535                 540

ttc gac aag acc atc aac aag ggc gac acc ctg acc tac aac agc ttc    1680
Phe Asp Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe
545                 550                 555                 560
```

```
aac ctg gcc agc ttc agc acc cct ttc gag ctg agc ggc aac aac ctc     1728
Asn Leu Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu
                565                 570                 575

cag atc ggc gtg acc ggc ctg agc gcc ggc gac aag gtg tac atc gac     1776
Gln Ile Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp
            580                 585                 590

aag atc gag ttc atc ccc gtg aac tag atc tga gct c                   1813
Lys Ile Glu Phe Ile Pro Val Asn     Ile     Ala
        595                 600
```

<210> 17
<211> 600
<212> PRT
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (322)..(333)
<223> Cathepsin G recognition site coding sequence

<220>
<221> misc_feature
<222> (1618)..(1629)
<223> Cathepsin G recognition site coding sequence

<400> 17

```
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1               5                   10                  15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
            20                  25                  30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
        35          .       40                  45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
    50                  55                  60

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

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
                85                  90                  95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Pro
            100                 105                 110

His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser His
        115                 120                 125
```

```
Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val Leu
    130                 135                 140

Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu Leu
145                 150                 155                 160

Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu Asp
                165                 170                 175

Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr Thr
            180                 185                 190

Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg Gly
        195                 200                 205

Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu Met
    210                 215                 220

Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp Val
225                 230                 235                 240

Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val Leu
                245                 250                 255

Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr Thr
            260                 265                 270

Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp Tyr
        275                 280                 285

Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr Gly
    290                 295                 300

Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg Pro
305                 310                 315                 320

Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn Lys
                325                 330                 335

Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val Tyr
            340                 345                 350

Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val Tyr
        355                 360                 365

Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr Asp
    370                 375                 380

Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala Val
385                 390                 395                 400

Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu Pro
                405                 410                 415
```

Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe Leu
420 425 430

Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His Lys
435 440 445

Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln Leu
450 455 460

Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val Ala
465 470 475 480

Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn Gly
485 490 495

Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln Lys
500 505 510

Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe Thr
515 520 525

Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Ala Ala Pro Phe Tyr
530 535 540

Phe Asp Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe
545 550 555 560

Asn Leu Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu
565 570 575

Gln Ile Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp
580 585 590

Lys Ile Glu Phe Ile Pro Val Asn
595 600

<210> 18
<211> 1819
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (1)..(1818)
<223> Maize optimized modified cry3A056 coding sequence.

<220>
<221> misc_feature
<222> (322)..(333)
<223> Catthepsin G recognition site coding sequence.

<220>
<221> misc_feature

<222> (1624)..(1635)
<223> Catthepsin G recognition site coding sequence.

<400> 18

```
atg acg gcc gac aac aac acc gag gcc ctg gac agc agc acc acc aag      48
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1               5                  10                  15

gac gtg atc cag aag ggc atc agc gtg gtg ggc gac ctg ctg ggc gtg      96
Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
            20                  25                  30

gtg ggc ttc ccc ttc ggc ggc gcc ctg gtg agc ttc tac acc aac ttc     144
Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
        35                  40                  45

ctg aac acc atc tgg ccc agc gag gac ccc tgg aag gcc ttc atg gag     192
Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
    50                  55                  60

cag gtg gag gcc ctg atg gac cag aag atc gcc gac tac gcc aag aac     240
Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65                  70                  75                  80

aag gca ctg gcc gag cta cag ggc ctc cag aac aac gtg gag gac tat     288
Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
                85                  90                  95

gtg agc gcc ctg agc agc tgg cag aag aac ccc gct gca ccg ttc cgc     336
Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Arg
            100                 105                 110

aac ccc cac agc cag ggc cgc atc cgc gag ctg ttc agc cag gcc gag     384
Asn Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu
        115                 120                 125

agc cac ttc cgc aac agc atg ccc agc ttc gcc atc agc ggc tac gag     432
Ser His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu
    130                 135                 140

gtg ctg ttc ctg acc acc tac gcc cag gcc gcc aac acc cac ctg ttc     480
Val Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe
145                 150                 155                 160

ctg ctg aag gac gcc caa atc tac gga gag gag tgg ggc tac gag aag     528
Leu Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys
                165                 170                 175

gag gac atc gcc gag ttc tac aag cgc cag ctg aag ctg acc cag gag     576
Glu Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu
            180                 185                 190

tac acc gac cac tgc gtg aag tgg tac aac gtg ggt cta gac aag ctc     624
Tyr Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu
```

```
        195                 200                 205

cgc ggc agc agc tac gag agc tgg gtg aac ttc aac cgc tac cgc cgc      672
Arg Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg
    210                 215                 220

gag atg acc ctg acc gtg ctg gac ctg atc gcc ctg ttc ccc ctg tac      720
Glu Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr
225                 230                 235                 240

gac gtg cgc ctg tac ccc aag gag gtg aag acc gag ctg acc cgc gac      768
Asp Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp
                245                 250                 255

gtg ctg acc gac ccc atc gtg ggc gtg aac aac ctg cgc ggc tac ggc      816
Val Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly
            260                 265                 270

acc acc ttc agc aac atc gag aac tac atc cgc aag ccc cac ctg ttc      864
Thr Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe
        275                 280                 285

gac tac ctg cac cgc atc cag ttc cac acg cgt ttc cag ccc ggc tac      912
Asp Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr
    290                 295                 300

tac ggc aac gac agc ttc aac tac tgg agc ggc aac tac gtg agc acc      960
Tyr Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr
305                 310                 315                 320

cgc ccc agc atc ggc agc aac gac atc atc acc agc ccc ttc tac ggc     1008
Arg Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly
                325                 330                 335

aac aag agc agc gag ccc gtg cag aac ctt gag ttc aac ggc gag aag     1056
Asn Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys
            340                 345                 350

gtg tac cgc gcc gtg gct aac acc aac ctg gcc gtg tgg ccc tct gca     1104
Val Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala
        355                 360                 365

gtg tac agc ggc gtg acc aag gtg gag ttc agc cag tac aac gac cag     1152
Val Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln
    370                 375                 380

acc gac gag gcc agc acc cag acc tac gac agc aag cgc aac gtg ggc     1200
Thr Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly
385                 390                 395                 400

gcc gtg agc tgg gac agc atc gac cag ctg ccc ccc gag acc acc gac     1248
Ala Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp
                405                 410                 415
```

```
gag ccc ctg gag aag ggc tac agc cac cag ctg aac tac gtg atg tgc      1296
Glu Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys
            420                 425                 430

ttc ctg atg cag ggc agc cgc ggc acc atc ccc gtg ctg acc tgg acc      1344
Phe Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr
        435                 440                 445

cac aag agc gtc gac ttc ttc aac atg atc gac agc aag aag atc acc      1392
His Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr
    450                 455                 460

cag ctg ccc ctg gtg aag gcc tac aag ctc cag agc ggc gcc agc gtg      1440
Gln Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val
465                 470                 475                 480

gtg gca ggc ccc cgc ttc acc ggc ggc gac atc atc cag tgc acc gag      1488
Val Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu
                485                 490                 495

aac ggc agc gcc gcc acc atc tac gtg acc ccc gac gtg agc tac agc      1536
Asn Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser
            500                 505                 510

cag aag tac cgc gcc cgc atc cac tac gcc agc acc agc cag atc acc      1584
Gln Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr
        515                 520                 525

ttc acc ctg agc ctg gac ggg gcc ccc ttc aac caa tac gct gca ccg      1632
Phe Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Ala Ala Pro
    530                 535                 540

ttc tac ttc gac aag acc atc aac aag ggc gac acc ctg acc tac aac      1680
Phe Tyr Phe Asp Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn
545                 550                 555                 560

agc ttc aac ctg gcc agc ttc agc acc cct ttc gag ctg agc ggc aac      1728
Ser Phe Asn Leu Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn
                565                 570                 575

aac ctc cag atc ggc gtg acc ggc ctg agc gcc ggc gac aag gtg tac      1776
Asn Leu Gln Ile Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr
            580                 585                 590

atc gac aag atc gag ttc atc ccc gtg aac tag atc tga gct c            1819
Ile Asp Lys Ile Glu Phe Ile Pro Val Asn     Ile     Ala
        595                 600
```

<210> 19
<211> 602
<212> PRT
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (322)..(333)
<223> Catthepsin G recognition site coding sequence.

<220>
<221> misc_feature
<222> (1624)..(1635)
<223> Catthepsin G recognition site coding sequence.

<400> 19

```
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1               5                  10                  15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
            20                  25                  30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
        35                  40                  45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
    50                  55                  60

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

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
                85                  90                  95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Arg
            100                 105                 110

Asn Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu
        115                 120                 125

Ser His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu
    130                 135                 140

Val Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe
145                 150                 155                 160

Leu Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys
                165                 170                 175

Glu Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu
            180                 185                 190

Tyr Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu
        195                 200                 205

Arg Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg
    210                 215                 220

Glu Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr
```

225 230 235 240

Asp Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp
245 250 255

Val Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly
260 265 270

Thr Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe
275 280 285

Asp Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr
290 295 300

Tyr Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr
305 310 315 320

Arg Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly
325 330 335

Asn Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys
340 345 350

Val Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala
355 360 365

Val Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln
370 375 380

Thr Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly
385 390 395 400

Ala Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp
405 410 415

Glu Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys
420 425 430

Phe Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr
435 440 445

His Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr
450 455 460

Gln Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val
465 470 475 480

Val Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu
485 490 495

Asn Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser
500 505 510

Gln Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr

515 520 525

Phe Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Ala Ala Pro
530 535 540

Phe Tyr Phe Asp Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn
545 550 555 560

Ser Phe Asn Leu Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn
565 570 575

Asn Leu Gln Ile Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr
580 585 590

Ile Asp Lys Ile Glu Phe Ile Pro Val Asn
595 600

<210> 20
<211> 1797
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (1)..(1791)
<223> Maize optimized modified cry3A083 coding sequence.

<220>
<221> misc_feature
<222> (322)..(333)
<223> Cathepsin G recognition site coding sequence.

<220>
<221> misc_feature
<222> (1612)..(1623)
<223> cathepsin G recognition site coding sequence

<400> 20
atg acg gcc gac aac aac acc gag gcc ctg gac agc agc acc acc aag 48
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

gac gtg atc cag aag ggc atc agc gtg gtg ggc gac ctg ctg ggc gtg 96
Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

gtg ggc ttc ccc ttc ggc ggc gcc ctg gtg agc ttc tac acc aac ttc 144
Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

ctg aac acc atc tgg ccc agc gag gac ccc tgg aag gcc ttc atg gag 192
Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

```
cag gtg gag gcc ctg atg gac cag aag atc gcc gac tac gcc aag aac      240
Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65                  70                  75                  80

aag gca ctg gcc gag cta cag ggc ctc cag aac aac gtg gag gac tat      288
Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
                85                  90                  95

gtg agc gcc ctg agc agc tgg cag aag aac ccc gct gca ccg ttc cgc      336
Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Arg
            100                 105                 110

aac ccc cac agc cag ggc cgc atc cgc gag ctg ttc agc cag gcc gag      384
Asn Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu
        115                 120                 125

agc cac ttc cgc aac agc atg ccc agc ttc gcc atc agc ggc tac gag      432
Ser His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu
    130                 135                 140

gtg ctg ttc ctg acc acc tac gcc cag gcc gcc aac acc cac ctg ttc      480
Val Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe
145                 150                 155                 160

ctg ctg aag gac gcc caa atc tac gga gag gag tgg ggc tac gag aag      528
Leu Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys
                165                 170                 175

gag gac atc gcc gag ttc tac aag cgc cag ctg aag ctg acc cag gag      576
Glu Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu
            180                 185                 190

tac acc gac cac tgc gtg aag tgg tac aac gtg ggt cta gac aag ctc      624
Tyr Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu
        195                 200                 205

cgc ggc agc agc tac gag agc tgg gtg aac ttc aac cgc tac cgc cgc      672
Arg Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg
    210                 215                 220

gag atg acc ctg acc gtg ctg gac ctg atc gcc ctg ttc ccc ctg tac      720
Glu Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr
225                 230                 235                 240

gac gtg cgc ctg tac ccc aag gag gtg aag acc gag ctg acc cgc gac      768
Asp Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp
                245                 250                 255

gtg ctg acc gac ccc atc gtg ggc gtg aac aac ctg cgc ggc tac ggc      816
Val Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly
            260                 265                 270

acc acc ttc agc aac atc gag aac tac atc cgc aag ccc cac ctg ttc      864
```

```
Thr Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe
        275                 280                 285

gac tac ctg cac cgc atc cag ttc cac acg cgt ttc cag ccc ggc tac      912
Asp Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr
    290                 295                 300

tac ggc aac gac agc ttc aac tac tgg agc ggc aac tac gtg agc acc      960
Tyr Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr
305                 310                 315                 320

cgc ccc agc atc ggc agc aac gac atc atc acc agc ccc ttc tac ggc     1008
Arg Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly
                325                 330                 335

aac aag agc agc gag ccc gtg cag aac ctt gag ttc aac ggc gag aag     1056
Asn Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys
            340                 345                 350

gtg tac cgc gcc gtg gct aac acc aac ctg gcc gtg tgg ccc tct gca     1104
Val Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala
        355                 360                 365

gtg tac agc ggc gtg acc aag gtg gag ttc agc cag tac aac gac cag     1152
Val Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln
    370                 375                 380

acc gac gag gcc agc acc cag acc tac gac agc aag cgc aac gtg ggc     1200
Thr Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly
385                 390                 395                 400

gcc gtg agc tgg gac agc atc gac cag ctg ccc ccc gag acc acc gac     1248
Ala Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp
                405                 410                 415

gag ccc ctg gag aag ggc tac agc cac cag ctg aac tac gtg atg tgc     1296
Glu Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys
            420                 425                 430

ttc ctg atg cag ggc agc cgc ggc acc atc ccc gtg ctg acc tgg acc     1344
Phe Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr
        435                 440                 445

cac aag agc gtc gac ttc ttc aac atg atc gac agc aag aag atc acc     1392
His Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr
    450                 455                 460

cag ctg ccc ctg gtg aag gcc tac aag ctc cag agc ggc gcc agc gtg     1440
Gln Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val
465                 470                 475                 480

gtg gca ggc ccc cgc ttc acc ggc ggc gac atc atc cag tgc acc gag     1488
Val Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu
                485                 490                 495
```

```
aac ggc agc gcc gcc acc atc tac gtg acc ccc gac gtg agc tac agc     1536
Asn Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser
            500                 505                 510

cag aag tac cgc gcc cgc atc cac tac gcc agc acc agc cag atc acc     1584
Gln Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr
        515                 520                 525

ttc acc ctg agc ctg gac ggg gcc ccc gct gca ccg ttc tac ttc gac     1632
Phe Thr Leu Ser Leu Asp Gly Ala Pro Ala Ala Pro Phe Tyr Phe Asp
    530                 535                 540

aag acc atc aac aag ggc gac acc ctg acc tac aac agc ttc aac ctg     1680
Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu
545                 550                 555                 560

gcc agc ttc agc acc cct ttc gag ctg agc ggc aac aac ctc cag atc     1728
Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile
                565                 570                 575

ggc gtg acc ggc ctg agc gcc ggc gac aag gtg tac atc gac aag atc     1776
Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile
            580                 585                 590

gag ttc atc ccc gtg aactag                                         1797
Glu Phe Ile Pro Val
        595
```

<210> 21
<211> 597
<212> PRT
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (322)..(333)
<223> Cathepsin G recognition site coding sequence.

<220>
<221> misc_feature
<222> (1612)..(1623)
<223> cathepsin G recognition site coding sequence

<400> 21

```
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1               5                   10                  15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
            20                  25                  30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
        35                  40                  45
```

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

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

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Arg
100 105 110

Asn Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu
115 120 125

Ser His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu
130 135 140

Val Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe
145 150 155 160

Leu Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys
165 170 175

Glu Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu
180 185 190

Tyr Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu
195 200 205

Arg Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg
210 215 220

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

Asp Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp
245 250 255

Val Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly
260 265 270

Thr Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe
275 280 285

Asp Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr
290 295 300

Tyr Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr
305 310 315 320

Arg Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly
325 330 335

Asn Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys
340 345 350

Val Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala
355 360 365

Val Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln
370 375 380

Thr Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly
385 390 395 400

Ala Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp
405 410 415

Glu Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys
420 425 430

Phe Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr
435 440 445

His Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr
450 455 460

Gln Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val
465 470 475 480

Val Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu
485 490 495

Asn Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser
500 505 510

Gln Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr
515 520 525

Phe Thr Leu Ser Leu Asp Gly Ala Pro Ala Ala Pro Phe Tyr Phe Asp
530 535 540

Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu
545 550 555 560

Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile
565 570 575

Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile
580 585 590

Glu Phe Ile Pro Val
595

<210> 22
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1)..(21)
<223> BamExt1 Primer

<400> 22
ggatccacca tgacggccga c 21

<210> 23
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1)..(29)
<223> AAPFtail3 Primer

<400> 23
gaacggtgca gcggggttct tctgccagc 29

<210> 24
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1)..(29)
<223> AAPFtail4 Primer

<400> 24
gctgcaccgt tcccccacag ccagggccg 29

<210> 25
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1)..(21)
<223> XbaIExt2 Primer

<400> 25
tctagaccca cgttgtacca c 21

<210> 26
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1)..(29)
<223> Tail5mod Primer

<400> 26
gctgcaccgt tccgcaaccc ccacagcca 29

<210> 27
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1)..(19)
<223> SalExt Primer

<400> 27
gagcgtcgac ttcttcaac 19

<210> 28
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1)..(30)
<223> AAPF-Y2 Primer

<400> 28
gaacggtgca gcgtattggt tgaaggggggc 30

<210> 29
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1)..(30)
<223> AAPF-Y1 Primer

<400> 29
gctgcaccgt tctacttcga caagaccatc 30

<210> 30
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1)..(21)
<223> SacExt Primer

<400> 30
gagctcagat ctagttcacg g 21

<210> 31
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1)..(32)
<223> BBmod1 Primer

<400> 31
cggggccccc gctgcaccgt tctacttcga ca 32

<210> 32
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1)..(32)
<223> BBmod2 Primer

<400> 32
tgtcgaagta gaacggtgca gcgggggccc cg 32

<210> 33
<211> 48
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature

<222> (1)..(48)
<223> mo3Aext Primer

<400> 33
ggatccacca tgaactacaa ggagttcctc cgcatgaccg ccgacaac 48

<210> 34
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<221> misc_feature
<222> (1)..(20)
<223> CMS16 Primer

<400> 34
cctccacctg ctccatgaag 20

Toxinas de Cry3A Modificadas y Secuencias de Acido Nucleico que las Codifican

La presente invención se refiere al campo de la manipulación genética de proteínas, a la biología molecular de plantas y al control de plagas. Más particularmente, la presente invención se refiere a nuevas toxinas Cry3A modificadas y a las secuencias de ácido nucleico cuya expresión da como resultado las toxinas Cry3A modificadas, y a los métodos para prepararlas y a los métodos de uso de las toxinas Cry3A modificadas y a las correspondientes secuencias de ácido nucleico para controlar insectos.

Las especies de gusanos de la raíz del maíz son considerados las plagas del maíz más destructivas. En los Estados Unidos, las tres especies importantes son *Diabrotica virgifera virgifera*, el gusano de la raíz del oeste; *D. longicornis barberi*, el gusano de la raíz del maíz del norte y *D. undecimpunctata howardi*, el gusano de la raíz del maíz del sur. Únicamente los gusanos de la raíz del maíz del oeste y del norte son considerados las plagas principales del maíz en el Cinturón del Maíz de los Estados Unidos. Las larvas del gusano de la raíz del maíz producen el daño más considerable en las plantas, debido a que se alimentan casi exclusivamente de las raíces del maíz. Este daño ha demostrado que aumenta el debilitamiento de las plantas, reduce el rendimiento de los granos y el rendimiento vegetativo, al mismo tiempo que altera el contenido de nutrientes del grano. La alimentación larval produce también efectos indirectos sobre el maíz al abrir avenidas a través de las raíces que fomentan las infecciones bacterianas y fungales, lo cual conduce a enfermedades tales como el pudrimiento del tallo y de la raíz. Los gusanos adultos de la raíz del maíz son activos en las postrimerías del verano en los maizales, en los cuales se alimentan de mazorcas, flores y polen, interfiriendo con la polinización normal.

Los gusanos de la raíz del maíz son controlados principalmente mediante aplicaciones intensivas de pesticidas químicos, que son activos a través de la

inhibición del desarrollo del insecto, impidiendo que el insecto se alimente o reproduzca, o causándole la muerte. Por lo tanto, debe obtenerse un buen control del gusano de la raíz del maíz, pero estos productos químicos pueden algunas veces afectar otros organismos que son beneficiosos. Otro problema resultante del amplio uso de los pesticidas químicos es la aparición de variedades resistentes a los insectos. Otro problema se debe al hecho de que las larvas del gusano de la raíz del maíz se alimentan subterráneamente, lo cual hace difícil aplicar tratamientos de rescate, con insecticidas. Por lo tanto, la mayoría de las aplicaciones insecticidas se efectúan profilácticamente en el momento de la siembra. Esta práctica da como resultado una gran carga ambiental. Esto se alivia parcialmente mediante varias prácticas de tratamiento en las fincas, pero existe una creciente necesidad de encontrar mecanismos alternativos para el control de plagas.

Agentes biológicos para el control de plagas, tales como las cepas de *Bacillus thuringiensis* (Bt) que expresan toxinas pesticidas tales como la endotoxina δ, han sido aplicados también con resultados satisfactorios contra plagas de insectos, principalmente lepidópteros. Las endotoxinas δ son proteínas mantenidas dentro de una matriz cristalina, las cuales se sabe que poseen actividad insecticida cuando son digeridos por ciertos insectos. Las varias endotoxinas δ han sido clasificadas en base a su espectro de actividad y homología de secuencia. Antes de 1990, las principales clases se definieron por su espectro de actividad con las proteínas Cry1 activas contra lepidópteros (polillas y mariposas), proteínas Cry2 proteínas activas tanto contyra lepidópteros como contra dípteros (moscas y mosquitos), proteínas Cry3 activas contra coleópteros (escarabajos) y proteínas Cry4 activas contra dípteros (Hofte and Whitely, 1989, Microbiol. Rev. 53:242-255). Recientemente, se ha desarrollado una nueva nomenclatura que clasifica sistemáticamente las proteínas Cry en base

a la homología de la secuencia de aminoácidos en vez de basarse en las especificidades del insecto que constituye el objetivo (Crickmore *et al.*, 1998, Microbiol, Molec. Biol. Rev. 62:807-813).

El espectro de la actividad insecticida de una endotoxina δ individual de Bt es bastante estrecho, ya que una determinada endotoxina δ es activa únicamente contra unas pocas especies dentro de un Orden. Por ejemplo, se sabe que la proteína Cry3A es muy tóxica para el escarabajo de la patata de Colorado, *Leptinotarsa decemlineata*, pero tiene muy poca o ninguna toxicidad con determinados escarabajos del género *Diabrotica* (Johnson *et al.*, 1993, J. Econ. Entomol. 86:330-333). De acuerdo con Slaney *et al* (1992, Insect Biochem. Molec. Biol. 22:9-18), la proteína Cry3A es por lo menos 2000 veces menos tóxica para las larvas del gusano de la raíz del maíz del sur que para el escarabajo de las patatas de Colorado. También es conocido que el Cry3A tiene poca o ninguna toxicidad para el gusano de la raíz del maíz del oeste.

La especificidad de las endotoxinas δ es el resultado de la eficiencia de las varias etapas que están involucradas en la producción de una proteína de toxina activa y su subsiguiente interacción con las células epiteliales en el mesenterón de los insectos. Para que sean insecticidas, la mayoría de las endotoxinas δ conocidas deben ser primero ingeridas por el insecto y activadas proteolíticamente para formar una toxina activa. La activación de las proteínas cristalinas insecticidas es un proceso de múltiples etapas. Después de la ingestión, los cristales deben ser primero solubilizados en el intestino del insecto. Una vez solubilizadas, las endotoxinas δ son activadas por disociaciones proteolíticas específicas. Las proteasas en el intestino del insecto pueden cumplir una función en lo que se refiere a la especificidad, mediante la determinación de donde se procesa la endotoxina δ. Una vez solubilizada y procesada la endotoxina δ, se une con los receptores específicos en la superficie del epitelio

del mesenterón de los insectos y subsiguientemente se integra en la bicapa de lípidos de la membrana de borde velloso. Entonces se forman los canales iónicos interrumpiendo la función normal del mesenterón y eventualmente llevando el insecto a la muerte.

En los lepidópteros, las proteasas intestinales procesan las endotoxinas δ de prototoxinas de 130-140 kDa a proteínas tóxicas de aproximadamente 60-70 kDa. Se ha descrito que el procesamiento de la prototoxina a toxina se efectúa mediante la remoción tanto de los aminoácidos N-terminales como los C-terminales, dependiendo la ubicación exacta del procesamiento, de los fluidos intestinales específicos del insecto involucrado (Ogiwara et al., 1992, J. Invert. Pathol. 60:121-126). La activación proteolítica de la endotoxina δ puede acumplir una función significativa en la determinación de su especificidad. Por ejemplo, una endotoxina δ de Bt var. *aizawa*, denominada ICI, ha sido clasificada como una proteína Cry1Ab en base a su homología de secuencia con otras proteínas Cry1Ab conocidas. Las proteínas Cry1Ab son típicamente activas contra insectos lepidópteros. Sin embargo, la proteína ICI tiene actividad contra insectos lepidópteros y dípteros dependiendo de la forma en la que se procesa la proteína (Haider et al., 1986, Euro. J. Biochem. 156:531-540). En el intestino de un díptero, se obtiene una toxina ICI activa de 53 kDa, mientras que en el intestino de un lepidóptero, se obtiene una toxina ICI activa de 55 kDa. La ICI difiere del holotipo HD-1 de la proteína Cry1Ab únicamente por cuatro aminoácidos, de modo que los cambios importantes en la región de unión del receptor no parecen contar para las diferencias de actividad. Las diferentes disociaciones proteolíticas en los dos intestinos de insectos diferentes, permiten posiblemente que las moléculas activadas se plieguen de manera diferente, lo cual expone diferentes regiones capaces de unirse a diferentes receptores. Por lo tanto la especificidad, parece residir en las proteasas del intestino de los diferentes insectos.

Los insectos coleópteros tienen intestinos que son más neutros a los ácidos y las endotoxinas δ específicas de coleópteros son similares al tamaño de las toxinas específicas de lepidópteros activadas. Por lo tanto, el procesamiento de las endotoxinas δ específicas de coleópteros se consideró en principio innecesario para la toxicidad. Sin embargo, datos recientes sugieren que las endotoxinas δ activas en coleópteros están solubilizadas y proteolizadas a polipéptidos tóxicos más pequeños. La proteína de las endotoxinas δ Cry3A de 73 kDa producida por *B. thuringiensis var. tenebrionis* es rápidamente procesada en la bacteria del término N, perdiendo los residuos 49-57 durante o después de la formación de los cristales, para producir la forma de 67 kDa comúnmente aislada (Carroll *et al.*, 1989, Biochem. J. 261:99-105). McPherson *et al.*, 1988 (Biotechnology 6:61-66) demostró también que el gen *cry*3A natural contiene dos codones de iniciación de transducción funcionales en el mismo marco de lectura, uno que codifica para la proteína de 63 kDa y otro que codifica para la proteína de 67 kDa a partir de Met-1 y Met-48 respectivamente, de la secuencia de aminoácido deducida (Ver SEC ID NO: 2). Ambas proteínas pueden considerarse entonces proteínas naturales Cry3A de longitud total. El tratamiento de la proteína soluble, Cry 3A de 67 kDa con tripsina o bien con extracto de intestino de insecto, da como resultado un producto de disociación de 55 kDa con el Asn-159 de la secuencia de aminoácido deducida en el término N. Se halló que este polipéptido era tan tóxico para un insecto coleóptero susceptible como la toxina Cry3A de 67 kDa (Carroll et al. *Ibid).* Por lo tanto, existe un sitio natural de reconocimiento de tripsina entre Arg-158 y Asn-159 de la secuencia de aminoácido deducida de la toxina natural Cry3A (SEC ID NO: 2). La Cry3A puede disociarse también con quimotripsina, lo cual da como resultado tres polipéptidos de 49, 11 y 6 kDa. El análisis N-terminal de los componentes de 49 y 6 kDa mostró que el primer residuo aminoácido era Ser-162 y Tyr-588, respectivamente (Carroll et al., 1997 J.

Invert. Biol. 70:41-49). Por lo tanto, existen sitios de reconocimiento de quimotripsina naturales en el Cry3A entre His-161 y Ser-162 y entre Tyr-587 y Tyr-588 de la secuencia de aminoácido deducida (SEC ID NO: 2). El producto de quimotripsina de 49 kDa parece ser más soluble a un pH neutro que la proteína de 67 kDa o el producto de tripsina de 55 kDa y retiene la actividad insecticida completa contra los insectos susceptibles a Cry3A, al escarabajo de la patata de Colorado, y al escarabajo de la mostaza *(Phaedon cochleariae)*. Las proteasas intestinales de insecto funcionan típicamente para asistir al insecto en la obtención de los aminoácidos necesarios de la proteína de su dieta. Las proteasas digestivas de los insectos, que se conocen mejor, son las serinas proteasas que parecen ser las más comunes (Englemann and Geraerts, 1980, J. Insect Physiol. 261:703-710), particularmente en las especies de lepidópteros. La mayoría de las larvas y adultos de coleópteros, por ejemplo del escarabajo de la patata de Colorado, tienen mesenterones ligeramente ácidos, y la cisteína proteasa provee la actividad proteolítica principal (Wolfson and Mudock, 1990, J. Chem. Ecol. 16.1089-1102). En forma más precisa, Thie y Houseman (1990, Insect. Biochem. 20:313-318) identificaron y caracterizaron las cisteínas proteasas, catepsina B y H, y la aspartil proteasa, catepsina D en el escarabajo de la patata de Colorado. Gillikin et al. (1992, Arch. Insect Biochem. Physiol. 19:285-298) caracterizó la actividad proteolítica en los intestinos de las larvas del gusano de la raíz del maíz del oeste y halló 15 proteasas, principalmente de cisteína. Hasta que se dio a conocer en esta invención, ninguno de los informes había indicado que la serina proteasa, catepsina G, existe en el gusano de la raíz del maíz del oeste. La diversidad y los diferentes niveles de actividad de las proteasas intestinales de insectos pueden influenciar la sensibilidad de un insecto a una toxina Bt particular.

Muchas y nuevas cepas de Bt y endotoxinas δ con actividades biológicas nuevas y mejoradas han sido descritas en los últimos cinco años, incluyendo las cepas que son activas contra los nematodos (EP 0517367A1). Sin embargo, relativamente muy pocas de estas cepas y toxinas tienen actividad contra los insectos coleópteros. Además, ninguna de las endotoxinas δ activas en coleópteros hasta el momento conocidas, por ejemplo, Cry3A, Cry3B, Cry3C, Cry7A, Cry8A, Cry8B, y Cry8C, tienen suficiente toxicidad oral contra el gusano de la raíz del maíz para proveer un control de campo adecuado, si se liberan, por ejemplo, a través de microbios o plantas transgénicas. Por lo tanto, es necesario explorar otras realizaciones para producir nuevas toxinas que puedan ser activas contra el gusano de la raíz del maíz.

Al aumentar el conocimiento en lo que se refiere a la forma en que funcionan las endotoxinas δ, se han incrementado los intentos de controlar las endotoxinas δ para nuevas actividades. La manipulación genética de las endotoxinas δ ha sido posible mediante la resolución de la estructura tridimensional del Cry3A en 1991 (Li *et al., 1991,* Nature 353:815-821). La proteína tiene tres dominios estructurales: el dominio I N-terminal de los residuos 1-290, que consiste en 7 hélices alfa, dominio II, de los residuos 291-500, que contiene tres láminas beta y el dominio C-terminal III, de los residuos 501-644, que es un sándwich beta. En base a esta estructura, se ha formulado la hipótesis con respecto a la relación de la estructura/función de las endotoxinas δ. Generalmente se cree que el dominio I es principalmente responsable de la formación de los poros en la membrana intestinal del insecto (Gazit and Shai, 1993, Appl. Environ. Microbiol. 57:2816-2820), el dominio II es principalmente responsable de la interacción con el receptor intestinal (Ge *et al.*, 1991, J. Biol. Chem. 32:3429-3436) y el dominio III muy probablemente está involucrado en la estabilidad de la proteína (Li *et al.*, 1991, supra) y tiene también un impacto

106
108

regulador sobre la actividad del canal iónico (Chen et al., 1993, PNAS 90:9041-9045).

Las endotoxinas δ activas en lepidópteros, han sido genéticamente manipuladas en un intento para mejorar la actividad específica o para ampliar el espectro de la actividad específica. Por ejemplo, el dominio de especificidad de la polilla de la seda *(Bómbix mori)* de Cry1Aa fue transferido a Cry1Ac, impartiendo de esta manera una nueva actividad insecticida a la proteína quimérica resultante (Ge *et al.*, 1989, PNAS 86:4037-4041). Asimismo, Bosch *et al*, 1998 Patente Norteamericana 5.736.131), creó una nueva toxina activa en lepidópteros sustituyendo el dominio III de Cry1E con el dominio III de Cry1C, produciendo de este modo una toxina híbrida Cry1E-Cry1C con un espectro de actividad más amplio en los lepidópteros.

Se ha informado sobre varios intentos para manipular genéticamente las endotoxinas δ en coleópteros. Van Rie *et al.* 1997, (Patente europea No. 5.659.123), el cual manipuló la Cry3A reemplazando al azar aminoácidos, creyó que era importante en la accesibilidad al solvente, en el dominio II con el aminoácido alanina. Varios de estos reemplazos al azar confinados al dominio II de unión al receptor, fueron repetidamente involucrados en la toxicidad incrementada del gusano de la raíz del maíz del oeste. Sin embargo, otros han demostrado que algunos reemplazos de alanina en el dominio II de Cry3A dieron como resultado la interrupción de la unión del receptor o una inestabilidad estructural (Wu and Dean, 1996, J. Mol. Biol. 255: 628-640). English et al., 1999 (Intl. Pat. Appl. Publ. No. WO 99/31248) informaron que las sustituciones de aminoácidos en Cry3Bb causaron aumentos de toxicidad en el gusano de la raíz del maíz del sur y del oeste. Sin embargo, de los 35 mutantes Cry3Bb informados, únicamente tres, con mutaciones principalmente en el dominio II y en la interfase del dominio II-dominio I, fueron activos contra el gusano de la raíz del maíz del

oeste. Además, las diferencias de toxicidad de Cry3Bb de tipo salvaje contra el gusano de la raíz del maíz del oeste en los mismos ensayos, fueron superiores a cualquiera de las diferencias entre las toxinas Cry3Bb mutadas y las Cry3Bb de tipo salvaje. Por lo tanto, las mejoras en la toxicidad de los mutantes Cry3Bb parecen estar principalmente confinadas al gusano de la raíz del maíz del sur.

Sigue existiendo la necesidad de diseñar nuevos y eficaces agentes de control para plagas que provean un beneficio económico para los agricultores y que sean ambientalmente aceptables. Particularmente se necesitan toxinas Cry3A modificadas que controlen al gusano de la raíz del maíz del oeste, que es la principal plaga del maíz en los Estados Unidos, y que son o podrían convertirse en resistentes a los agentes de control de insectos existentes. Además, son convenientes agentes cuya aplicación minimice la carga ambiental, a través de plantas transgénicas.

En vista de estas necesidades, es un objeto de la presente invención proveer nuevas secuencias de ácido nucleico que codifiquen para toxinas Cry3A modificadas, que tengan toxicidad incrementada para el gusano de la raíz del maíz. Mediante la inserción de un sitio de reconocimiento de la proteasa, que es reconocido por una proteasa intestinal del insecto-objetivo en por lo menos una posición de una toxina Cry3A de acuerdo con la presente invención, se diseñó una toxina Cry3A modificada que tiene significativamente mayor toxicidad, particularmente para el gusano de la raíz del maíz del oeste y del norte. La invención está además dirigida a las nuevas toxinas Cry3A modificadas, resultantes de la expresión de la secuencia de ácido nucleico, y a las composiciones y formulaciones que contienen las toxinas Cry3A modificadas, las cuales son capaces de inhibir la capacidad de los insectos para sobrevivir, crecer y reproducirse o son capaces de limitar el perjuicio relacionado con los insectos o la pérdida de plantas de cultivo. La invención está además dirigida a un método

para preparar las toxinas Cry3A modificadas y a los métodos de uso de la secuencia de ácido nucleico de Cry3A, por ejemplo en los microorganismos para controlar insectos y plantas transgénicas, para conferir protección contra el perjuicio causado por los insectos, y a un método para usar las toxinas Cry3A modificadas, y las composiciones y formulaciones que comprenden las toxinas Cry3A modificadas, por ejemplo, para aplicar las toxinas Cry3A modificadas o las composiciones o formulaciones a las áreas infestadas por insectos, o para tratar profilácticamente las áreas o plantas susceptibles a los insectos y conferirles protección contra las plagas de insectos.

Las nuevas toxinas Cry3A modificadas aquí descritas son sumamente activas contra los insectos. Por ejemplo, las toxinas Cry3A modificadas de la presente invención, puede usarse para controlar plagas de insectos económicamente importantes tales como el gusano de la raíz del maíz del oeste *(Diabrotica virgifera virgifera)* y el gusano de la raíz del maíz del norte (*D. longicornis barberi)*. Las toxinas Cry3A modificadas pueden usarse solas o en combinación con otras estrategias para controlar insectos con el fin de conferirles una máxima eficiencia para el control de plagas con un mínimo impacto ambiental.

De acuerdo con un aspecto, la presente invención provee una molécula de ácido nucleico aislada que comprende una secuencia nucleotídica que codifica una toxina Cry3A modificada, donde la toxina Cry3A modificada comprende por lo menos un sitio adicional de reconocimiento de proteasa que no se encuentra en forma natural en una toxina Cry3A. El sitio de reconocimiento de proteasa adicional, que es reconocido por una proteasa intestinal de un insecto objetivo, se inserta aproximadamente en la misma posición que un sitio de reconocimiento de proteasa natural, en la toxina Cry3A. La toxina Cry3A modificada causa una mortalidad superior para un insecto objetivo que la mortalidad causada por una

toxina Cry3A en el mismo insecto objetivo. Preferiblemente, la toxina Cry3A modificada causa por lo menos aproximadamente 50% de mortalidad para un insecto objetivo en el cual la toxina Cry3A causa únicamente hasta aproximadamente 30% de mortalidad.

En una realización de este aspecto, la proteasa intestinal de un insecto objetivo se selecciona del grupo que consiste en las serina proteasas, las cisteína proteasas y las proteasas aspárticas. Preferiblemente, las serina proteasas de acuerdo con esta realización incluyen catepsina G, tripsina, quimotripsina, carboxipeptidasa, endopeptidasa y elastasa, más preferiblemente catepsina G.

En otra realización de este aspecto, el sitio de reconocimiento adicional de la proteasa se inserta o en el dominio I o bien en el dominio III o en ambos dominios, el dominio I y el III de las toxinas Cry3A.

Preferiblemente, el sitio de reconocimiento de la proteasa adicional se inserta en cualquiera de los dominio I o dominio III o en ambos, el dominio I y el dominio III, en una posición que reemplaza, que es adyacente, o que está dentro de un sitio de reconocimiento de proteasa natural.

En una realización más, el sitio de reconocimiento de proteasa adicional se inserta en el dominio I entre los aminoácidos que corresponden a los aminoácidos números 154 y 162 de la SEC ID NO:2. Preferiblemente, el sitio de reconocimiento de la proteasa adicional se inserta entre los números de aminoácidos 154 y 162 de la SEC ID NO. 2 o entre los números de aminoácidos 107 y 115 de la SEC ID NO. 4.

En otra realización más, el sitio de reconocimiento de proteasa adicional se inserta en el dominio III entre los aminoácidos que corresponden a los aminoácidos números 154 y 160 de la SEC ID NO: 2. Preferiblemente, el sitio de reconocimiento de proteasa adicional entre los aminoácidos número 154 y 160 de la SEC ID NO: 2 o entre los aminoácidos números 107 y 113 de la SEC ID NO: 4.

En una realización más, el sitio de reconocimiento de proteasa adicional se inserta en el dominio I entre los aminoácidos números 154 y 158 de la SEC ID NO: 2. Preferiblemente, el sitio de reconocimiento de proteasa adicional se inserta en el dominio I entre los aminoácidos número 154 y 158 de la SEC ID NO: 2 o entre los aminoácidos número 107 y 111 de la SEC ID NO: 4.

En otra realización, el sitio de reconocimiento de proteasa adicional se inserta en el dominio III entre los aminoácidos que corresponden a los aminoácidos números 583 y 589 de la SEC ID NO: 2. Preferiblemente, el sitio de proteasa adicional se inserta en el dominio III entre los aminoácidos números 583 y 589 de la SEC ID NO: 2, o entre los aminoácidos números 536 y 542 de la SEC ID NO: 4.

En otra realización más, el sitio de reconocimiento de proteasa adicional se inserta en el dominio III entre los aminoácidos que corresponden a los aminoácidos números 583 y 588 de la SEC ID NO: 2. Preferiblemente, el sitio de proteasa adicional se inserta en el dominio III entre los aminoácidos números 583 y 588 de la SEC ID NO: 2 o entre los aminoácidos números 536 y 541 de la SEC ID NO: 4.

En otra realización, el sitio de reconocimiento de proteasa adicional se inserta en el dominio III entre los aminoácidos que corresponden a los aminoácidos número 587 y 588 de la SEC ID NO: 2. Preferiblemente, el sitio de proteasa adicional se inserta en el dominio III entre los aminoácidos números 587 y 588 de la SEC ID NO: 2 o entre los aminoácidos número 540 y 541 de la SEC ID NO: 4.

En otra realización, el sitio de reconocimiento de proteasa adicional se inserta en el dominio I y en el dominio III de las toxinas Cry3A no modificada. Preferiblemente, el sitio de reconocimiento de proteasa adicional se inserta en el dominio I, en una posición que reemplaza o que es adyacente a un sitio de

reconocimiento de proteasa natural, y en el dominio III en una posición que está dentro, que reemplaza, o que es adyacente a un sitio de reconocimiento de proteasa natural.

En otra realización, el sitio de reconocimiento de proteasa adicional se inserta en el dominio I o entre los aminoácidos que corresponden a los aminoácidos número 154 y 160, y en el dominio III entre los aminoácidos que corresponden a los aminoácidos números 587 y 588 de la SEC ID NO: 2. Preferiblemente, el sitio de reconocimiento de proteasa adicional se inserta en el dominio I entre los aminoácidos números 154 y 160 y en el dominio III entre los aminoácidos números 587 y 588 de la SEC ID NO: 2 o en el dominio I entre los aminoácidos números 107 y 113 y en el dominio III entre los aminoácidos números 540 y 541 de la SEC ID NO: 4.

En otra realización, el sitio de reconocimiento de proteasa adicional está situado en el dominio I entre los aminoácidos que corresponden a los aminoácidos números 154 y 158 y en el dominio III entre los aminoácidos que corresponden a los aminoácidos números 587 y 588 de la SEC ID NO: 2. Preferiblemente, el sitio de reconocimiento de proteasa adicional se inserta en el dominio I entre los aminoácidos números 154 y 158 y en el dominio III entre los aminoácidos números 587 y 588 de la SEC ID NO: 2 o en el dominio I entre los aminoácidos números 107 y 111 y en el dominio III entre los aminoácidos números 540 y 541 de la SEC ID NO: 4.

En otra realización, el sitio de reconocimiento de proteasa adicional está situado en el dominio I entre los aminoácidos que corresponden a los aminoácidos números 154 y 158 y en el dominio III entre los aminoácidos que corresponden a los aminoácidos números 583 y 588 de la SEC ID NO: 2. Preferiblemente, el sitio de reconocimiento de proteasa adicional se inserta en el dominio I entre los aminoácidos números 154 y 158 y en el dominio III entre los

aminoácidos números 583 y 588 de la SEC ID NO: 2 o en el dominio I entre los aminoácidos números 107 y 111 y en el dominio III entre los aminoácidos números 536 y 541 de la SEC ID NO: 4.

En una realización preferida, la molécula de ácido nucleico aislada de la presente invención comprende los nucleótidos 1-1791 de la SEC ID NO: 6, los nucleótidos 1-1806 de la SEC ID NO: 8, los nucleótidos 1-1818 de la SEC ID NO: 10, los nucleótidos 1-1794 de la SEC ID NO: 12, los nucleótidos 1-1812 de la SEC ID NO: 14, los nucleótidos 1-1812 de la SEC ID NO: 16, los nucleótidos 1-1818 de la SEC ID NO: 18 o los nucleótidos 1-1791 de la SEC ID NO: 20.

En otra realización preferida, la molécula de ácido nucleico aislado de la invención codifica para la toxina Cry3A modificada que comprende la secuencia de aminoácido establecida en las SEC ID NO: 7, SEC ID NO: 9, SEC ID NO: 11, SEC ID NO: 13, SEC ID NO: 15, SEC ID NO: 17, SEC ID NO: 19, SEC ID NO: 21.

De acuerdo con una realización de la invención, la molécula de ácido nucleico aislada codifica una toxina Cry3A modificada que es activa contra un insecto coleóptero. Preferiblemente, la toxina Cry3A modificada tiene actividad contra el gusano de la raíz del maíz del oeste.

La presente invención provee un gen quimérico que comprende una secuencia promotora heteróloga ligada operativamente a la molécula de ácido nucleico de la invención. La presente invención provee también un vector recombinante que comprende dicho gen quimérico. Además, la presente invención provee una célula huésped no humana transgénica que comprende dicho gen quimérico. Una célula huésped transgénica de acuerdo con este aspecto de la invención puede ser una célula bacteriana o una célula de planta, preferiblemente una célula de planta. La presente invención provee además una planta transgénica que comprende dicha célula de planta. Una planta transgénica de acuerdo con este aspecto de la invención puede ser sorgo, trigo, girasol,

tomate, patata, cultivos de coles, algodón, arroz, soja, remolacha, caña de azúcar, tabaco, cebada, colza de semilla oleosa o maíz, preferiblemente, maíz. La presente invención provee también semillas del grupo de las plantas transgénicas que consisten en sorgo, trigo, girasol, tomate, patata, cultivo de coles, algodón, arroz, soja, remolacha, caña de azúcar, tabaco, cebada, colza de semilla oleosa y maíz. En una realización particularmente preferida, la semilla procede de una planta de maíz transgénica. En otro aspecto, la presente invención provee toxinas producidas mediante la expresión de las moléculas de ácido nucleico de la presente invención. En una realización preferida, la toxina se produce mediante la expresión de la molécula de ácido nucleico que comprende los nucleótidos 1-1791 de la SEC ID NO: 6, los nucleótidos 1-1806 de la SEC ID NO: 8, los nucleótidos 1-1818 de la SEC ID NO: 10, los nucleótidos 1-1794 de la SEC ID NO: 12, los nucleótidos 1-1812 de la SEC ID NO: 14, los nucleótidos 1-1812 de la SEC ID NO: 16, los nucleótidos 1-1818 de la SEC ID NO: 18 o los nucleótidos 1-1791 de la SEC ID NO: 20.

En otra realización, las toxinas de la invención son activas contra insectos coleópteros, preferiblemente contra el gusano de la raíz del maíz del oeste.

En otra realización, una toxina de la presente invención comprende la secuencia de aminoácidos establecida en las SEC ID NO: 7, SEC ID NO: 9, SEC ID NO: 11, SEC ID NO: 13, SEC ID NO: 15, SEC ID NO: 17, SEC ID NO: 19, SEC ID NO: 21.

La presente invención provee también una composición que comprende una cantidad eficaz para el control de insectos, de una toxina de acuerdo con la invención.

En otro aspecto, la presente invención provee un método para producir una toxina que es activa contra insectos, que comprende: (a) obtener una célula huésped que comprende un gen quimérico, la cual comprende por sí misma una

secuencia promotora heteróloga ligada operativamente a la molécula de ácido nucleico de la invención; y (b) expresar la molécula de ácido nucleico en la célula huésped transgénica, lo cual da como resultado por lo menos una toxina que es activa contra insectos.

En otro aspecto, la presente invención provee un método para producir una planta transgénica resistente a los insectos, que comprende introducir una molécula de ácido nucleico de la invención, en la planta transgénica, donde la molécula de ácido nucleico se expresa en la planta transgénica, en una cantidad eficaz para controlar los insectos. En una realización preferida, los insectos son insectos coleópteros, preferiblemente gusanos de la raíz del maíz del oeste.

En otro aspecto adicional, la presente invención provee un método para controlar insectos, que comprende liberar a los insectos una cantidad eficaz de una toxina de la invención. De acuerdo con una realización, los insectos son insectos coleópteros, preferiblemente el gusano de la raíz del maíz del oeste. Preferiblemente, la toxina es liberada a los insectos por vía oral. En una realización preferida, la toxina es liberada oralmente a través de una planta transgénica que comprende una secuencia de ácido nucleico que expresa una toxina de la presente invención.

Asimismo, se provee mediante la presente invención un método para preparar una toxina Cry3A modificada que comprende: (a) obtener un gen de toxina *cry*3A que codifica una toxina Cry3A; (b) identificar una proteasa intestinal de un insecto objetivo; (c) obtener una secuencia nucleotídica que codifica una secuencia de reconocimiento para la proteasa intestinal; (d) insertar la secuencia nucleotídifica de (c) en el dominio I o bien en el dominio III o en ambos, el dominio I y el dominio III, en una posición que reemplaza, que está dentro o que es adyacente a una secuencia nucleotídica que codifica para un sitio de reconocimiento de proteasa que ocurre en forma natural en un gen de una toxina

*cry*3A creando de esta manera un gen de toxina *cry*3A modificado; (e) insertar el gen de la toxina *cry*3A modificado en un cassette de expresión; (f) expresar el gen de toxina Cry3A modificado en una célula huésped no humana, lo cual da como resultado que la célula huésped produzca una toxina Cry3A modificada; y, (g) bioensayar la toxina Cry3A modificada contra un insecto objetivo, mediante lo cual la toxina Cry3A modificada produce una mortalidad superior para el insecto objetivo, que la mortalidad causada por una toxina Cry3A. En una realización preferida, la toxina Cry3A modificada causa por lo menos aproximadamente un 50% de mortalidad para el insecto objetivo cuando la toxina Cry3A causa hasta aproximadamente 30% de mortalidad. La presente invención provee además un método para controlar insectos en el cual la planta transgénica comprende además, una segunda secuencia de ácido nucleico o un grupo de secuencias de ácido nucleico que codifican un segundo principio pesticida. Segundas secuencias de ácido nucleico particularmente preferidas son aquellas que codifican una endotoxina δ, aquellas que codifican una toxina de Proteína Insecticida Vegetativa, descritas en las Patentes Norteamericanas Nº 5.849.870 y 5.877.012, que se incorporan aquí como referencia, o las que codifican una vía para la producción de un principio pesticida no proteináceo.

Un aspecto más de la presente invención es la provisión de un método para mutagenizar una molécula de ácido nucleico de acuerdo con la presente invención, donde la molécula de ácido nucleico ha sido disociada dentro de poblaciones de fragmentos al azar de doble hélice, de un tamaño deseado, que comprende: (a) agregar a la población de los fragmentos al azar de doble hélice, uno o más oligonucleótidos de un solo filamento o de doble hélice, donde los oligonucleótidos comprenden cada uno un área de identidad y un área de heterología para un patrón polinucleótido de doble hélice; (b) desnaturalizar la mezcla resultante de fragmentos y oligonucleótidos al azar de doble hélice en

fragmentos de un solo filamento, (c) incubar la población resultante de los fragmentos de un solo filamento con polimerasa bajo condiciones que dan como resultado el acoplamiento de los fragmentos de un solo filamento en las áreas de identidad para formar pares de fragmentos acoplados, siendo las áreas de identidad suficientes para que un miembro del par cebe la replicación del otro, formando de este modo un polinucleótido de doble hélice mutagenizado; y (d) repetir la segunda y tercera etapas durante por lo menos dos ciclos más, donde la mezcla resultante en la segunda etapa de un ciclo adicional incluye el polinucleótido de doble hélice mutagenizado de la tercera etapa del ciclo previo, y donde el ciclo adicional forma otro polinucleótido de doble hélice mutagenizado.

Otros aspectos y ventajas de la presente invención resultarán aparentes para los expertos en la materia a partir de un estudio de la siguiente descripción de la invención y de los ejemplos no limitativos.

SEC ID NO: 1 es la región codificadora *cry*3A natural.

SEC ID NO: 2 es la secuencia de aminoácido de la toxina Cry3A codificada por el gen *cry*3A natural.

SEC ID NO: 3 es la región codificadora *cry*3A optimizada de maíz, que comienza en el nucleótido 144 de la región codificadora de *cry*3A natural.

SEC ID NO: 4 es la secuencia de aminoácido de la toxina Cry3A codificada por el gen *cry*3A optimizado del maíz.

SEC ID NO: 5 es la secuencia nucleotídica de pCIB6850.

SEC ID NO: 6 es la secuencia codificadora de *cry*3A054 modificada optimizada de maíz.

SEC ID NO: 7 es la secuencia de aminoácido codificada por la secuencia nucleotídica de la SEC ID NO: 6.

SEC ID NO: 8 es la secuencia codificadora *cry*3A055 modificada optimizada de maíz.

SEC ID NO: 9 es la secuencia de aminoácido codificada por la secuencia nucleotídica de la SEC ID NO: 8.

SEC ID NO: 10 es la secuencia codificadora *cry*3A085 modificada optimizada de maíz.

SEC ID NO: 11 es la secuencia de aminoácido codificada por la secuencia nucleotídica de la SEC ID NO: 10.

SEC ID NO: 12 es la secuencia codificadora *cry*3A082 modificada optimizada de maíz.

SEC ID NO: 13 es la secuencia de aminoácido codificada por la secuencia nucleotídica de la SEC ID NO: 12.

SEC ID NO: 14 es la secuencia codificadora *cry*3A058 modificada optimizada de maíz.

SEC ID NO: 15 es la secuencia de aminoácido codificada por la secuencia nucleotídica de la SEC ID NO: 14.

SEC ID NO: 16 es la secuencia codificadora *cry*3A057 modificada optimizada de maíz.

SEC ID NO: 17 es la secuencia de aminoácido codificada por la secuencia nucleotídica de la SEC ID NO: 16.

SEC ID NO: 18 es la secuencia codificadora *cry*3A056 modificada optimizada de maíz.

SEC ID NO: 19 es la secuencia de aminoácido codificada por la secuencia nucleotídica de la SEC ID NO: 18.

SEC ID NO: 20 es la secuencia codificadora *cry*3A083 modificada optimizada de maíz.

SEC ID NO: 21 es la secuencia de aminoácido codificada por la secuencia nucleotídica de la SEC ID NO: 20.

SEC ID NOS: 22-34 son los cebadores de PCR útiles en la presente invención.

SEC ID NO: 35: es una secuencia de aminoácido que comprende un sitio de reconocimiento de la catepsina G.

SEC ID NO: 36: es una secuencia de aminoácido que comprende un sitio de reconocimiento de la catepsina G.

SEC ID NO: 37: es una secuencia de aminoácido que comprende un sitio de reconocimiento de la catepsina G.

SEC ID NO: 38: es una secuencia de aminoácido que comprende un sitio de reconocimiento de la catepsina G.

Por razones de claridad, algunos términos usados en la memoria se definen y presentan a continuación: "Actividad" de las toxinas Cry3A modificadas de la invención significa que las toxinas Cry3A funcionan como agentes de control de insectos oralmente activos, que tienen un efecto tóxico, o que son capaces de interrumpir o impedir la alimentación del insecto, lo cual puede causar o no la muerte del insecto. Cuando una toxina Cry3A modificada de la invención es liberada al insecto, el resultado es típicamente la muerte del insecto, o el insecto no se alimenta de la fuente que produce la toxina Cry3A modificada asequible para el insecto.

"Adyacente a" – De acuerdo con la presente invención, un sitio de reconocimiento de proteasa es "adyacente a" un sitio de reconocimiento de proteasa natural cuando el sitio de reconocimiento de proteasa adicional está dentro de cuatro residuos, preferiblemente dentro de tres residuos, más preferiblemente dentro de dos residuos, y aún más preferiblemente dentro de un residuo de un sitio de reconocimiento de proteasa que ocurre en forma natural. Por ejemplo, un sitio de reconocimiento de proteasa adicional insertado entre Pro-

154 y Arg-158 de la secuencia de aminoácido deducida de una toxina Cry3A (SEC ID NO: 2), es "adyacente" al sitio de reconocimiento de tripsina natural situado entre Arg-158 y Asn-159 de la secuencia de aminoácido deducida de la toxina Cry3A (SEC ID NO: 2).

La frase "aproximadamente la misma posición" tal como se usa aquí para describir el lugar del sitio de reconocimiento de proteasa adicional en una toxina Cry3A en relación a un sitio de reconocimiento de proteasa natural, significa que el lugar está a lo sumo cuatro residuos alejado de un sitio de reconocimiento de proteasa natural. El lugar puede estar también tres o dos residuos alejado de un sitio de reconocimiento de proteasa natural. El lugar puede estar también un residuo alejado de un sitio de reconocimiento de proteasa natural. "Aproximadamente la misma posición" significa también que el sitio de reconocimiento de proteasa adicional está insertado dentro de un sitio de reconocimiento de proteasa natural.

"Asociado con / ligado operativamente" se refiere a dos secuencias de ácido nucleico que están física o funcionalmente relacionadas. Por ejemplo, se dice que un promotor o una secuencia de ADN reguladora está "asociada con" una secuencia de ADN que codifica para una proteína de ARN, si las dos secuencias están ligadas operativamente, o están situadas de manera que la secuencia reguladora de ADN afectará el nivel de expresión de la secuencia codificadora o estructural de ADN .

Un "gen quimérico" o "construcción quimérica" es una secuencia de ácido nucleico recombinante en la cual una secuencia de ácido nucleico promotora o reguladora está ligada operativamente o está asociada con una secuencia de ácido nucleico que codifica para un ARNm o que se expresa como proteína, de manera que la secuencia de ácido nucleico reguladora será capaz de regular la transcripción y expresión de la secuencia codificadora de ácido nucleico asociada.

La secuencia de ácido nucleico reguladora del gen quimérico no está normalmente ligada operativamente a la secuencia de ácido nucleico asociada tal como se encuentra en la naturaleza.

Una "secuencia codificadora" es una secuencia de ácido nucleico que está transcripta en el ARN, tal como ARNm, ARNr, ARNt, ARNsn, ARN de sentido o ARN de anti-sentido. Preferiblemente el ARN es luego traducido en un organismo para producir una proteína.

"Controlar" insectos significa inhibir, a través de un efecto tóxico, la capacidad de las plagas de insectos para sobrevivir, crecer, alimentarse y/o reproducirse, o para limitar el daño relacionado con los insectos o la pérdida de plantas de cultivo. "Controlar" insectos puede significar o no la muerte de los insectos, aunque preferiblemente significa la muerte de los insectos.

Correspondiente a: en el contexto de la presente invención, "correspondiente a" significa que cuando la secuencias de aminoácidos de las endotoxinas δ Cry3A variables están alineadas una con la otra, los aminoácidos que "corresponden" a ciertas posiciones enumeradas en la presente invención, son aquellos que se alinean con estas posiciones en la toxina Cry3A (SEC ID NO: 2), pero que no están necesariamente en estas posiciones numéricas exactas en relación a la secuencia de aminoácido cry3A particular de la invención. Por ejemplo, el gen *cry*3A optimizado de maíz (SEC ID NO:3) de la invención codifica una toxina Cry3A (SEC ID NO: 4) que comienza en Met-48 de la toxina Cry3A (SEC ID NO: 2) codificada por el gen *cry*3A natural (SEC ID NO: 1). Por lo tanto, de acuerdo con la presente invención, los números de aminoácidos 107-115, incluyendo todos los números comprendidos entre lós mismos, y 536-541, incluyendo todos los números comprendidos entre los mismos, de la SEC ID NO: 4 corresponden a los aminoácidos números 154-163, y todos los números entre

los mismos y 583-588, y todos los números entre los mismos, respectivamente, de la SEC ID NO: 2.

Una "toxina Cry3A", tal como se usa aquí, se refiere a una proteína activa en coleópteros (Bt) de *Bacillus thuringiensis var. tenebrionis* de aproximadamente 73 kDa (Kreig *et al.*, 1983, Z. Angew. Entomol. 96:500-508) (Sekar *et al.*, 1987, Proc. Natl. Acad. Sci. 84:7036-7040), por ejemplo la SEC ID NO: 2, así como también cualquier variante truncada de menor peso molecular, derivable de una toxina Cry3A, por ejemplo, la SEC ID NO: 4, y que retiene sustancialmente la misma toxicidad que las toxinas Cry3A. Pueden obtenerse variantes de peso molecular inferior mediante la disociación de proteasa en los sitios de reconocimiento de proteasa natural de las toxinas Cry3A, o mediante un segundo codón de iniciación de traducción en el mismo marco que el codón de iniciación de traducción que codifica para la toxina Cry3A de 73 kDa. La secuencia de aminoácido de una toxina Cry3A y sus variantes de peso molecular inferior, pueden encontrarse en una toxina natural que ocurre en Bt. Una toxina Cry3A puede estar codificada por un gen Bt natural tal como la SEC ID NO: 1, o por una secuencia codificadora sintética tal como la SEC ID NO: 3. Una "toxina Cry3A" no tiene ningún sitio de reconocimiento de proteasa adicional sobre los sitios de reconocimiento de proteasa que ocurren en forma natural en la toxina Cry3A. Una toxina Cry3A puede ser aislada, purificada o expresada en un sistema heterólogo.

Un "gen *cry*3A" tal como se usa aquí, se refiere a la secuencia nucleotídica de la SEC ID NO: 1 ó SEC ID NO: 3. Un gen *cry*3A (Sekar *et al.*, 1987, Proc. Natl. Acad. Sci. 84:7036-7040) puede ser natural, tal como el que se encuentra en el *Bacillus thuringiensis var. tenebrionis* (Kreig *et al.*, 1983, Z. Angew. Entomol. 96:500-508), o puede ser sintético y codificar una toxina Cry3A. El gen *cry*3A de esta invención puede denominarse gen *cry*3A natural tal como en la SEC ID NO: 1, o gen *cry*3A optimizado de maíz tal como en la SEC ID NO: 3.

"Liberar" una toxina significa que la toxina entra en contacto con un insecto, lo cual da como resultado un efecto tóxico y el control del insecto. La toxina puede ser liberada en muchas formas conocidas, por ejemplo oralmente al ser ingerida por el insecto, o por contacto con el insecto a través de la expresión de una planta transgénica, composición(es) de proteína formulada(s), composición(es) de proteína pulverizable(s), una matriz de cebo, o cualquier otro sistema de liberación de toxinas reconocido en el arte.

"Cantidad eficaz para el control de insectos" se refiere a la concentración de toxina que inhibe, a través de un efecto tóxico, la capacidad de los insectos para sobrevivir, crecer, alimentarse y/o reproducirse, o para limitar el daño relacionado con los insectos o la pérdida de plantas de cultivo. "Cantidad eficaz para el control de insectos" puede o no significar la muerte de los insectos, aunque preferiblemente significa la muerte de los insectos.

"Cassette de expresión" tal como se usa aquí se refiere a una secuencia de ácido nucleico que es capaz de dirigir la expresión de una secuencia nucleotídica particular en una célula huésped apropiada, que comprende un promotor ligado operativamente a la secuencia nucleotídica de interés, la cual está operativamente ligada a las señales de terminación. Asimismo, típicamente, comprende la traducción adecuada de la secuencia nucleotídica. El cassette de expresión que comprende la secuencia nucleotídica de interés puede ser quimérico, lo cual significa que por lo menos uno de sus componentes es heterólogo con respecto a por lo menos uno de sus otros componentes. El cassette de expresión puede ser un cassette natural pero que ha sido obtenido en una forma recombinante útil para la expresión heteróloga. Típicamente, sin embargo, el cassette de expresión es heterólogo con respecto al huésped, es decir, la secuencia de ácido nucleico particular del cassette de expresión no ocurre en forma natural en la célula huésped y debe ser introducida dentro de la

célula huésped o un ancestro de la célula huésped mediante un evento de transformación. La expresión de la secuencia nucleotídica en el cassette de expresión puede estar bajo el control de un promotor constitutivo o de un promotor inducible que inicia la transcripción únicamente cuando la célula huésped es expuesta a algún estímulo externo particular. En el caso de un organismo multicelular, tal como una planta, el promotor puede ser también un promotor específico para un residuo u órgano en la etapa de desarrollo particular.

Un "gen" es una región definida que está situada dentro de un genoma y que, además de la secuencia de ácido nucleico codificadora antes mencionada, comprende otras secuencias de ácido nucleico, principalmente reguladoras, que son responsables del control de la expresión, es decir, de la transcripción y traducción, de la porción codificadora. Un gen puede comprender también otras secuencias no traducidas 5' y 3' y secuencias de terminación. Otros elementos que pueden estar presentes son por ejemplo intrones.

"Gen de interés" se refiere a cualquier gen que, cuando es transferido a una planta, confiere a la planta una característica deseada tal como resistencia a un antibiótico, resistencia a un virus, resistencia a los insectos, resistencia a las enfermedades, o resistencia a otras plagas, tolerancia a los herbicidas, mejor valor nutricional, mejor rendimiento en un proceso industrial o capacidad reproductora alterada. El "gen de interés" puede ser también un gen que es transferido a plantas para la producción de enzimas o metabolitos comercialmente valiosos en la planta.

"Una proteasa intestinal" es una proteasa natural que se encuentra en el tracto digestivo de un insecto. Esta proteasa está usualmente involucrada en la digestión de las proteínas ingeridas.

Una secuencia de ácido nucleico "heteróloga" es una secuencia de ácido nucleico que no está naturalmente asociada con una célula hospedante dentro de

la cual es introducida, incluyendo múltiples copias que no ocurren en forma natural, de una secuencia de ácido nucleico que ocurre en forma natural.

Una secuencia de ácido nucleico "homóloga" es una secuencia de ácido nucleico que está naturalmente asociada con una célula hospedante dentro de la cual es introducida.

"Recombinación homóloga" es el intercambio reciproco de fermentos de ácido nucleico entre moléculas de ácido nucleico homólogas.

"Insecticida" se define como una actividad biológica tóxica que es capaz de controlar insectos, preferiblemente matándolos.

Una secuencia de ácido nucleico está "isocodificada con" una secuencia de ácido nucleico de referencia cuando la secuencia de ácido nucleico codifica un polipéptido que tiene la misma secuencia de aminoácido que el polipéptido codificado por la secuencia de ácido nucleico de referencia.

Una molécula de ácido nucleico "aislada" o una toxina aislada es una molécula de ácido nucleico o una toxina que, producida por la mano del hombre, existe aparte de su ambiente natural y que por lo tanto no es un producto de la naturaleza. Una molécula o toxina de ácido nucleico aislada puede existir en una forma purificada o puede existir en un ambiente no natural, tal como por ejemplo una célula huésped recombinante.

Una "toxina Cry3A modificada" de esta invención, se refiere a una toxina derivada de Cry3A que tiene por lo menos un sitio de reconocimiento de proteasa adicional que es reconocido por una proteasa intestinal de un insecto objetivo, que no ocurre en forma natural en una toxina Cry3A. Una toxina Cry3A modificada no es una toxina natural, y es producida por la mano del hombre y comprende una secuencia de aminoácido que no es idéntica a la de una toxina natural que se encuentra en el *Bacillus thuringiensis*. La toxina Cry3A modificada produce una

mortalidad más elevada en un insecto objetivo que la mortalidad causada por una toxina Cry3A para el mismo insecto objetivo.

Un "gen *cry*3A modificado" de acuerdo con esta invención, se refiere a un gen derivado de *cry*3A que comprende la secuencia codificadora de por lo menos un sitio de reconocimiento de proteasa adicional que no ocurre en forma natural en un gen *cry*3A no modificado. El gen *cry*3A modificado puede estar derivado de un gen *cry*3A natural o de un gen *cry*3A sintético.

Un "sitio de reconocimiento de proteasa natural" es un lugar en el cual se disocia una toxina Cry3A mediante una proteasa que no deriva de insectos, o mediante una proteasa o extracto intestinal de una especie de insecto que es susceptible a la toxina Cry3A. Por ejemplo, un sitio de reconocimiento de una proteasa natural, reconocido por tripsina y por proteasas que se encuentran en un extracto intestinal de insecto susceptible, existe entre Arg-158 y Asn-159 de la secuencia de aminoácido de la toxina Cry3A deducida (SEC ID NO: 2). Los sitios de reconocimiento de proteasas naturales reconocidos por quimotripsina, existen entre His-161 y Ser-162 así como también entre Tyr-587 y Tyr-588 de la secuencia de aminoácido de la toxina Cry3A deducida (SEC ID NO: 2).

Una "molécula de ácido nucleico" o "secuencia de ácido nucleico" es un segmento lineal de un ARN o ADN de un solo filamento o de doble hélice, que puede aislarse de cualquier fuente. En el contexto de la presente invención, la molécula de ácido nucleico es preferiblemente un segmento de ADN.

Una "planta" es cualquier planta en cualquier etapa de desarrollo, particularmente una planta de semilla.

Una "célula de planta" es una unidad estructural y fisiológica de una planta, que comprende un protoplasto y una pared celular. La célula de plantas puede estar en forma de una sola célula aislada, o una célula cultivada, o puede formar

parte de una unidad organizada, tal como por ejemplo, tejido de plantas, órgano de una planta o una planta entera.

"Cultivo de células de planta" significa cultivos de unidades de plantas, tales como por ejemplo protoplastos, células de cultivo de células, células de tejidos de plantas, polen, tubos de polen, óvulos, bolsas embriónicas, cigotos y embriones en varias etapas de desarrollo.

"Material de planta" se refiere a las hojas, tallos, raíces, flores o partes de flores, frutos, polen, células de óvulos, cigotos, semillas, cortes, cultivos de células o de tejido o cualquier otra parte de un producto de una planta.

"Organo de una planta" es una parte distintiva y visiblemente estructurada y diferenciada de una planta, tal como una raíz, un tallo, una hoja, un brote de una flor o un embrión.

"Tejido de planta" tal como se usa aquí significa un grupo de células de plantas organizado en una unidad estructural y funcional. Se incluye cualquier tejido de una planta *in planta* o de cultivo. Este término incluye pero no está limitado a, plantas enteras, órganos de plantas, semillas de plantas, cultivos de tejidos y cualquier grupo de células de plantas organizadas en unidades estructurales y/o funcionales. El uso de este término conjuntamente o en ausencia de cualquier tipo específico de tejido de planta enumerado anteriormente o en otra parte, abarcado por esta definición, no pretende ser exclusivo de cualquier otro tipo de tejido de planta.

Un "promotor" es una secuencia de ADN no traducida cadena arriba de la región codificadora, que contiene el sitio de unión para la polimerasa de ARN y que inicia la transcripción del ADN. La región promotora puede incluir también otros elementos que actúan como reguladores de la expresión génica.

Un "protoplasto" es una célula de planta aislada sin una pared celular o con únicamente partes de la pared celular.

"Elementos reguladores" se refiere a las secuencias que están involucradas en el control de la expresión de una secuencia nucleotídica. Los elementos reguladores comprenden un promotor ligado operativamente a la secuencia nucleotídica de interés y señales de terminación. Abarcan también típicamente las secuencias requeridas para una adecuada traducción de la secuencia nucleotídica.

"Reemplaza" un sitio de reconocimiento de proteasa natural - De acuerdo con la presente invención, un sitio de reconocimiento de proteasa adicional "reemplaza" un sitio de reconocimiento de proteasa natural cuando la inserción del sitio de reconocimiento de proteasa adicional elimina el sitio de reconocimiento de proteasa natural. Por ejemplo, un sitio de reconocimiento de proteasa adicional insertado entre Pro-154 y Pro-160 de la secuencia de aminoácido deducida de una toxina Cry3A (SEC ID NO: 2) que elimina los residuos Arg-158 y Asn-159 "reemplaza" el sitio de reconocimiento de tripsina natural situado entre Arg-158 y Asn-159 de la secuencia de aminoácido deducida de la toxina Cry3A (SEC ID NO: 2).

"Serina proteasas" describe el mismo grupo de enzimas que catalizan la hidrólisis de enlaces peptídicos covalentes usando un mecanismo basado en el ataque nucleofílico del enlace peptídico que constituye el objetivo mediante una serina. Las serina proteasas son específicas de secuencia. Es decir, cada una de las serina proteasas reconoce una sub-secuencia específica dentro de una proteína en la cual ocurre un reconocimiento enzimático.

Un "insecto objetivo" es una especie de plaga de insectos que tiene poca o ninguna susceptibilidad para las toxinas Cry3A y que se identifica como un candidato para usar la tecnología de la presente invención como control. Este control puede lograrse a través de varios medios, pero más preferiblemente a

través de la expresión de las moléculas de ácido nucleico de la invención en plantas transgénicas.

Una "proteasa intestinal de insecto objetivo" es una proteasa que se encuentra en el intestino de un insecto objetivo cuyo sitio de reconocimiento puede insertarse en una toxina Cry3A para crear una toxina Cry3A modificada de la invención.

"Transformación es un procedimiento para introducir un ácido nucleico heterólogo dentro de una célula u organismo huésped. En particular, "transformación" significa la integración estable de una molécula de ADN dentro del genoma de un organismo de interés.

"Transformado / transgénico / recombinante" se refiere a un organismo huésped tal como una bacteria o una planta dentro de la cual se ha introducido una molécula de ácido nucleico heteróloga. La molécula de ácido nucleico puede estar establemente integrada dentro del genoma del huésped o puede estar también presente en la molécula de ácido nucleico como molécula extra cromosomal. Dicha molécula extracromosomal puede ser auto-replicante. Se entiende que células, tejidos o plantas transformadas abarca no solamente el producto final de un proceso de transformación, sino que también abarca la progenie del mismo. Un huésped "no transformado", "no transgénico" o "no recombinante" se refiere a un organismo de tipo salvaje, por ejemplo, una bacteria o planta, que no contiene la molécula de ácido nucleico heteróloga.

"Dentro" de un sitio de reconocimiento de proteasa natural - De acuerdo con la presente invención, un sitio de reconocimiento de proteasa adicional está "dentro" de un sitio de reconocimiento de proteasa natural cuando el sitio de reconocimiento de proteasa adicional está entre el residuo de aminoácido que viene antes y el residuo de aminoácido que viene después del sitio de reconocimiento de la proteasa natural. Por ejemplo, un sitio de reconocimiento de

proteasa adicional insertado entre Tyr-587 y Tyr-588 de la secuencia de aminoácido deducida de una toxina Cry3A (SEC ID NO: 2) está "dentro de un sitio de reconocimiento de quimotripsina natural situado entre Tyr-587 y Tyr-588 de la secuencia de aminoácido deducida de las toxinas Cry3A (SEC ID NO: 2). La inserción de un sitio de reconocimiento de proteasa adicional dentro de un sitio de reconocimiento de proteasa natural puede o no cambiar el reconocimiento del sitio de reconocimiento de proteasa natural mediante una proteasa.

Los nucleotidos están indicados por sus bases mediante las siguientes abreviaturas convencionales: adenina (A), citosina (C), timina (T), y guanina (G). Los aminoácidos están asimismo indicados mediante las siguientes abreviaturas convencionales: alanina (Ala; A), arginina (Arg; R), asparagina (Asn; N), ácido aspártico (Asp; D), cisteína (Cys; C), glutamina (Gln; Q), ácido glutámico (Glu; E), glicina (Gly; G), histidina (His; H), isoleucina (Ile; 1), leucina (Leu; L), lisina (Lys; K), metionina (Met; M), fenilalanina (Phe; F), prolina (Pro; P), serina (Ser; S), treonina (Thr; T), triptofano (Trp; W), tirosina (Tyr; Y), y valina (Val; V).

Esta invención se refiere a secuencias de ácido nucleico *cry*3A modificadas cuya expresión da como resultado toxinas Cry3A modificadas, y a la preparación y uso de las toxinas *cry*3A modificadas para controlar plagas de insectos. La expresión de las secuencias de ácido nucleico *cry*3A modificadas da como resultado toxinas Cry3A modificadas que pueden usarse para controlar insectos coleópteros tales como el gusano de la raíz del maíz del oeste y el gusano de la raíz del maíz del norte. Una toxina Cry3A modificada de la presente invención comprende por lo menos un sitio de reconocimiento de proteasa adicional que no ocurre naturalmente en una toxina Cry3A. El sitio de reconocimiento de proteasa adicional que es reconocido por una proteasa intestinal de un insecto objetivo, se inserta aproximadamente en la misma posición que un sitio de reconocimiento de proteasa natural en una toxina Cry3A. La toxina Cry3A modificada produce una

mortalidad más elevada en un insecto objetivo que la mortalidad causada por una toxina Cry3A para el mismo insecto objetivo. Preferiblemente la toxina Cry3A modificada causa por lo menos aproximadamente un 50% de mortalidad en el insecto objetivo en el cual una toxina Cry3A causa hasta aproximadamente un 30% de mortalidad.

En una realización preferida, la invención abarca una molécula de ácido nucleico aislada que codifica una toxina Cry3A modificada, donde el sitio de reconocimiento de proteasa adicional es reconocido por una proteasa intestinal de insecto objetivo, la catepsina G. Se ha determinado que la actividad de la catepsina G está presente en el intestino del insecto objetivo, el gusano de la raíz del oeste, tal como se ha descrito en el Ejemplo 2. Preferiblemente, la secuencia de aminoácido de sustrato AAPF (SEC ID NO: 35) usada para determinar la presencia de la actividad de la catepsina G, se inserta dentro de la toxina Cry3A de acuerdo con la presente invención. Otros sitios de reconocimiento de catepsina G pueden usarse también de acuerdo con la presente invención, por ejemplo, AAPM (SEC ID NO: 36), AVPF (SEC ID NO: 37), PFLF (SEC ID NO: 38) u otros sitios de reconocimiento de catepsina G de acuerdo con lo determinado por el método de Tanaka *et al.*, 1985 (Biochemistry 24: 2040-2047), que se incorpora aquí como referencia. Pueden usarse otros sitios de reconocimiento de proteasa de otras proteasas identificadas en un intestino de insecto objetivo, por ejemplo los sitios de reconocimiento de proteasa reconocidos por otras serina proteasas, cisteína proteasas, y proteasas aspárticas. Las serina proteasas preferidas que están abarcadas en esta realización, incluyen tripsina, quimotripsina, carboxipeptidasa, endopeptidasa y elastasa.

En otra realización preferida, la invención abarca una molécula de ácido nucleico aislada que codifica una toxina Cry3A modificada donde el sitio de reconocimiento de proteasa adicional está insertado en el dominio I o en el

dominio III o en ambos, el dominio I y dominio III de la toxina Cry3A. Preferiblemente, el sitio de reconocimiento de la proteasa adicional se inserta en el dominio I ó dominio III, o en el dominio I y dominio III en una posición que reemplaza, que es adyacente o que está dentro de un sitio de reconocimiento de la proteasa natural en las toxinas Cry3A. Se ejemplifican aquí específicamente las moléculas de ácido nucleico que codifican las toxinas Cry3A modificadas que comprenden un sitio de reconocimiento de la catepsina G insertado en el dominio I o dominio III o en el dominio I y dominio III en una posición que reemplaza, y que es adyacente o que está dentro de un sitio de reconocimiento de proteasa natural, en las toxinas Cry3A no modificadas.

En el Ejemplo 3 pueden encontrarse las enseñanzas específicamente ejemplificadas de los métodos para preparar moléculas de ácido nucleico *cry*3A que codifican toxinas Cry3A modificadas. Los expertos en la materia reconocerán que pueden usarse también otros métodos conocidos en el arte para insertar sitios de reconocimiento de proteasa adicionales, en las toxinas Cry3A de acuerdo con la presente invención. En otra realización preferida, la invención abarca una molécula de ácido nucleico aislada que codifica una toxina Cry3A modificada donde el sitio de reconocimiento de proteasa adicional se inserta en el dominio I entre los aminoácidos que corresponden a los aminoácidos números: 154 y 162 de la SEC ID NO: 2. Preferiblemente el sitio de reconocimiento de proteasa adicional se inserta entre los aminoácidos números: 154 y 162 de la SEC ID NO: 2 ó entre los aminoácidos números: 107 y 115 de la SEC ID NO: 4. En una realización preferida, el sitio de reconocimiento de proteasa adicional se inserta entre los aminoácidos que corresponden los aminoácidos números: 154 y 160 de la SEC ID NO: 2. Preferiblemente, el sitio de reconocimiento de proteasa adicional se inserta entre los aminoácidos números: 154 y 160 de la SEC ID NO: 2 o entre los aminoácidos números: 107 y 113 de la SEC ID NO: 4. Se ejemplifica aquí

específicamente, una molécula de ácido nucleico, designada *cry*3A054 (SEC ID NO: 6), que codifica la toxina Cry3A054 modificada (SEC ID NO: 7) que comprende un sitio de reconocimiento de catepsina G insertado en el dominio I entre los aminoácidos números: 107 y 113 de la SEC ID NO: 4. El sitio de reconocimiento de catepsina G reemplaza un sitio de reconocimiento de tripsina natural que es adyacente a un sitio de reconocimiento de quimotripsina natural. Cuando se expresa en un huésped heterólogo, la molécula de ácido nucleico de la SEC ID NO: 6 da como resultado una actividad para el control de insectos contra el gusano de la raíz del maíz del oeste y contra el gusano de la raíz del maíz del norte, lo cual muestra que la secuencia de ácido nucleico establecida en la SEC ID NO: 6 es suficiente para dicha actividad para el control de insectos.

En otra realización preferida, el sitio de reconocimiento de proteasa adicional se inserta en el dominio I entre los aminoácidos que corresponden a los aminoácidos números 154 y 158 de la SEC ID NO: 2. Preferiblemente, el sitio de reconocimiento de proteasa adicional se inserta en el dominio I entre los aminoácidos números 154 y 158 de la SEC ID NO: 2 o entre los aminoácidos números 107 y 111 de la SEC ID NO: 4. Se ejemplifican aquí específicamente las moléculas de ácido nucleico designadas *cry*3A055 (SEC ID NO: 8), que codifican la toxina Cry3A055 modificada (SEC ID NO: 9) y *cry*3A085 (SEC ID NO: 10), que codifica la toxina Cry3A085 (SEC ID NO: 11), que comprende un sitio de reconocimiento de catepsina G insertado en el dominio I entre los aminoácidos números 107 y 111 de la SEC ID NO: 4. El sitio de reconocimiento de la catepsina G es adyacente a los sitios de reconocimiento de tripsina y quimotripsina naturales. Cuando se expresa en un huésped heterólogo, la molécula de ácido nucleico de la SEC ID NO: 8 y de la SEC ID NO: 10 da como resultado una actividad para el control de insectos contra el gusano de la raíz del maíz del oeste y el gusano de la raíz del maíz del norte, lo cual muestra la secuencia de ácido

nucleico establecida en la SEC ID NO: 8 o la SEC ID NO: 10 es suficiente para dicha actividad para el control de insectos. En una realización preferida, la invención abarca una molécula de ácido nucleico aislada que codifica una toxina Cry3A modificada donde el sitio de reconocimiento de proteasa adicional se inserta en el dominio III entre los aminoácidos que corresponden a los aminoácidos números 583 y 589 de la SEC ID NO: 2. Preferiblemente el sitio de proteasa adicional se inserta en el dominio III entre los aminoácidos números 583 y 589 de la SEC ID NO: 2 o entre los aminoácidos números 536 y 542 de la SEC ID NO: 4.

En otra realización preferida, la invención abarca una molécula de ácido nucleico aislada que codifica una toxina Cry3A modificada donde el sitio de reconocimiento de proteasa adicional está insertado en el dominio III entre los aminoácidos que corresponden a los aminoácidos números 583 y 588 de la SEC ID NO: 2. Preferiblemente, el sitio de proteasa adicional está insertado en el dominio III entre los aminoácidos números 583 y 588 de la SEC ID NO: 2 o entre los aminoácidos números 536 y 541 de la SEC ID NO: 4. Se ejemplifica específicamente aquí una molécula de ácido nucleico designada *cry*3A082 (SEC ID NO: 12) que codifica la toxina Cry3A082 modificada (SEC ID NO: 13) que comprende un sitio de reconocimiento de catepsina G insertado en el dominio III entre los aminoácidos números 536 y 541 de la SEC ID NO: 4. El sitio de reconocimiento de catepsina G reemplaza un sitio de reconocimiento de quimotripsina natural. Cuando se expresa en un huésped heterólogo, la molécula de ácido nucleico de la SEC ID NO: 12 da como resultado una actividad para el control de insectos contra el gusano de la raíz del maíz del oeste y contra el gusano de la raíz del maíz del norte, que muestra que la secuencia de ácido nucleico establecida en la SEC ID NO: 12 es suficiente para dicha actividad de control de insectos.

En otra realización preferida, el sitio de reconocimiento de proteasa adicional está insertado en el dominio III entre los aminoácidos que corresponden a los aminoácidos números 587 y 588 de la SEC ID NO: 2. Preferiblemente, el sitio de proteasa adicional está insertado en el dominio III entre los aminoácidos números 587 y 588 de la SEC ID NO: 2 o entre los aminoácidos números 540 y 541 de la SEC ID NO: 4. Se ejemplifica específicamente aquí una molécula de ácido nucleico designada *cry*3A058 (SEC ID NO: 14), que codifica una toxina Cry3A058 modificada (SEC ID NO: 15), que comprende un sitio de reconocimiento de catepsina G insertado en el dominio III entre los aminoácidos números 540 y 541 de la SEC ID NO. 4. El sitio de reconocimiento de catepsina G está dentro de un sitio de reconocimiento de quimotripsina natural. Cuando se expresa en un huésped heterólogo, la molécula de ácido nucleico de la SEC ID NO: 14 da como resultado una actividad para el control de insectos contra el gusano de la raíz del maíz del oeste y contra el gusano de la raíz del maíz del norte, que muestra que la secuencia de ácido nucleico establecida en la SEC ID NO: 14 es suficiente para dicha actividad para el control de insectos.

En una realización preferida, la invención abarca una molécula de ácido nucleico aislada que codifica una toxina Cry3A modificada donde el sitio de reconocimiento de proteasa adicional está insertado en el dominio I entre los aminoácidos que corresponden a los aminoácidos número 154 y 160 y en el dominio III entre los aminoácidos que corresponden a los aminoácidos números 587 y 588 de la SEC ID NO: 2. Preferiblemente, el sitio de reconocimiento de proteasa adicional está insertado en el dominio I entre los aminoácidos números 154 y 160 y en el dominio III entre los aminoácidos números 587 y 588 de la SEC ID NO: 2 o en el dominio I entre los aminoácidos números 107 y 113 y en el dominio III entre los aminoácidos números 540 y 541 de la SEC ID NO: 4.

Se ejemplifica específicamente aquí una molécula de ácido nucleico, designada *cry*3A057 (SEC ID NO: 16), que codifica la toxina Cry3A057 modificada (SEC ID NO: 17) que comprende un sitio de reconocimiento de catepsina G insertado en el dominio I entre los aminoácidos números 107 y 113 y en el dominio III entre los aminoácidos números 540 y 541 de la SEC ID NO: 4. el sitio de reconocimiento de catepsina G reemplaza un sitio de reconocimiento de tripsina natural y es adyacente a un sitio de reconocimiento de quimotripsina natural en el dominio I y está dentro de un sitio de un reconocimiento de quimotripsina natural en el dominio III. Cuando se expresa en un huésped heterólogo, la molécula de ácido nucleico de la SEC ID NO: 16 da como resultado una actividad para el control de insectos contra el gusano de la raíz del maíz del oeste y contra el gusano de la raíz del maíz del norte, lo cual muestra que la secuencia de ácido nucleico establecida en la SEC ID NO: 16 es suficiente para dicha actividad para el control de insectos.

En otra realización preferida, el sitio de reconocimiento de proteasa adicional está situado en el dominio I entre los aminoácidos que corresponden a los aminoácidos números 154 y 158 y en el dominio III entre los aminoácidos que corresponden a los aminoácidos números 587 y 588 de la SEC ID NO: 2. Preferiblemente el sitio de reconocimiento de proteasa adicional está insertado en el dominio I entre los aminoácidos números 154 y 158 y en el dominio III entre los aminoácidos números 587 y 588 de la SEC ID NO: 2 o en el dominio I entre los aminoácidos números 107 y 111 y en el dominio III entre los aminoácidos números 540 y 541 de la SEC ID NO: 4. Ejemplificada en forma específica aquí está la molécula de ácido nucleico designada *cry*3A056 (SEC ID NO: 18), que identifica la toxina Cry3A056 modificada (SEC ID NO: 19) que comprende un sitio de reconocimiento de catepsina G insertado en el dominio I entre los aminoácidos números 107 y 111 y en el dominio III entre los aminoácidos números 540 y 541

de la SEC ID NO: 4. El sitio de reconocimiento de la catepsina G es adyacente a los sitios de reconocimiento de tripsina y quimotripsina naturales en el dominio I y está dentro del sitio de reconocimiento de quimotripsina natural en el dominio III. Cuando se expresa en un huésped heterólogo, la molécula de ácido nucleico de la SEC ID NO: 18 da como resultado una actividad para el control de insectos contra el gusano de la raíz del maíz del oeste y el gusano de la raíz del maíz del norte, que muestra que la secuencia de ácido nucleico establecida en la SEC ID NO: 18 es suficiente para dicha actividad para el control de insectos. En otra realización preferida, el sitio de reconocimiento de proteasa adicional está situado en el dominio I entre los aminoácidos que corresponden a los aminoácidos número 154 y 158 y en el dominio III entre los aminoácidos que corresponden a los aminoácidos números 583 y 588 de la SEC ID NO: 2. Preferiblemente, el sitio de reconocimiento de la proteasa adicional está insertado en el dominio I entre los aminoácidos números 154 y 158 y en el dominio III entre los aminoácidos números 583 y 588 de la SEC ID NO: 2 o en el dominio I entre los aminoácidos número 107 y 111 y en el dominio III entre los aminoácidos números 536 y 541 de la SEC ID NO: 4. Está ejemplificada específicamente aquí, una molécula de ácido nucleico, designada *cry*3A083 (SEC ID NO: 20), que codifica la toxina Cry3A083 modificada (SEC ID NO: 21) que comprende un sitio de reconocimiento de catepsina G insertada en el dominio I entre los aminoácidos números 107 y 111 y en el dominio III entre los aminoácidos números 536 y 541 de la SEC ID NO: 4. El sitio de reconocimiento de la catepsina G es adyacente a los sitios de reconocimiento de tripsina y quimotripsina naturales en el dominio I y reemplaza un sitio de reconocimiento de quimotripsina natural en el dominio III. Cuando se expresa en un huésped heterólogo, la molécula de ácido nucleico de la SEC ID NO: 20 da como resultado una actividad para el control de insectos contra el gusano de la raíz del maíz del oeste y el gusano de la raíz del maíz del norte, que

muestra que la secuencia de ácido nucleico establecida en la SEC ID NO: 20 es suficiente para dicha actividad para el control de insectos. En una realización preferida, la molécula de ácido nucleico aislada de la presente invención comprende los nucleótidos 1-1791 de la SEC ID NO: 6, los nucleótidos 1-1806 de la SEC ID NO: 8, los nucleótidos 1-1818 de la SEC ID NO: 10, los nucleótidos 1-1794 de la SEC ID NO: 12, los nucleótidos 1-1812 de la SEC ID NO: 14, los nucleótidos 1-1812 de la SEC ID NO: 16, los nucleótidos 1-1791 de la SEC ID NO: 18 o los nucleótidos 1-1818 de la SEC ID NO: 20.

En otra realización preferida, la invención abarca la molécula de ácido nucleico aislada que codifica una toxina Cry3A modificada que comprende la secuencia de aminoácido establecida en las SEC ID NO: 7, SEC ID NO: 9, SEC ID NO: 11, SEC ID NO: 13, SEC ID NO: 15, SEC ID NO: 17, SEC ID NO: 19, SEC ID NO: 21.

La presente invención abarca también los vectores recombinantes que comprenden la secuencia de ácido nucleico de esta invención. En dichos vectores, las secuencias de ácido nucleicos están preferiblemente comprendidas en los cassettes de expresión que comprenden los elementos reguladores para la expresión de las secuencias nucleotídicas en una célula huésped capaz de expresar las secuencias nucleotídicas. Dichos elementos reguladores comprenden usualmente señales promotoras y de terminación y preferiblemente comprenden también elementos que permiten una eficiente traducción de los polipéptidos codificados por la secuencia de ácido nucleico de la presente invención. Los vectores que comprenden las secuencias de ácido nucleico son usualmente capaces de replicación en las células huésped particulares, preferiblemente, moléculas extracromosomales y por lo tanto se usan para amplificar las secuencias de ácido nucleico de esta invención en las células huésped. En una realización, las células huésped de dichos vectores son

microorganismos, tales como bacterias, en particular, *Bacillus thuringiensis* o *E. coli*. En otra realización, las células huésped para dichos vectores recombinantes son endofitos o epifitos. Una célula huésped preferida para dichos vectores es una célula eucariótica tal como una célula de planta. Las células de plantas tales como células de maíz son las células huésped preferidas. En otra realización preferida, dichos vectores son vectores virales y se usan para la replicación de las secuencias nucleotídicas, en particular de células huésped, por ejemplo, células de insectos o células de plantas. También se usan vectores recombinantes para la transformación de las secuencias nucleotídicas de esta invención en células huésped, mediante lo cual las secuencias nucleotídicas quedan integradas establemente dentro del ADN de dichas células huésped. En una realización dichas células huésped son células procarióticas. En una realización preferida, dichas células huésped son células eucarióticas, tales como células de plantas. En una realización aún más preferida, las células huésped son células de planta tales como células de maíz. En otro aspecto, la presente invención abarca toxinas Cry3A modificadas producidas por la expresión de las moléculas de ácido nucleico de la presente invención.

En realizaciones preferidas, las toxinas Cry3A modificadas de la invención comprenden un polipéptido codificado por una secuencia nucleotídica de la invención. En una realización más preferida, la toxina Cry3A modificada se produce mediante la expresión de la molécula de ácido nucleico que comprende los nucleótidos 1-1791 de la SEC ID NO: 6, los nucleótidos 1-1806 de la SEC ID NO: 8, los nucleótidos 1-1812 de la SEC ID NO: 10, los nucleótidos 1-1794 de la SEC ID NO: 12, los nucleótidos 1-1818 de la SEC ID NO: 14, los nucleótidos 1-1812 de la SEC ID NO: 16, los nucleótidos 1-1791 de la SEC ID NO: 18 o los nucleótidos 1-1818 de la SEC ID NO: 20.

En una realización preferida, una toxina Cry3A modificada de la presente invención comprende la secuencia de aminoácido establecida en las SEC ID NO: 7, SEC ID NO: 9, SEC ID NO: 11, SEC ID NO: 13, SEC ID NO: 15, SEC ID NO: 17, SEC ID NO: 19, SEC ID NO: 21.

Las toxinas Cry3A modificadas de la presente invención tienen actividad para el control de insectos cuando son probadas contra plagas de insectos en bioensayos. En otra realización preferida, las toxinas Cry3A modificadas de la invención son activas contra insectos coleópteros, preferiblemente contra el gusano de la raíz del maíz del oeste y el gusano de la raíz del maíz del norte. Las propiedades controladoras de insectos de las toxinas Cry3A modificadas de la invención se ilustran adicionalmente en los ejemplos 4 y 6.

La presente invención abarca también una composición que comprende una cantidad eficaz controladora de insectos de una toxina Cry3A modificada de la invención. En otra realización preferida, la invención abarca un método para producir una toxina Cry3A modificada que es activa contra insectos, que comprende (a) obtener una célula huésped que comprende un gen quimérico, el cual comprende por sí mismo una secuencia promotora heteróloga ligada operativamente a la molécula de ácido nucleico de la invención; y (b) que expresa la molécula de ácido nucleico en la célula huésped transgénica, lo cual da como resultado por lo menos una toxina Cry3A modificada que es activa contra insectos.

En otra realización preferida, la invención abarca un método para producir una planta transgénica resistente a los insectos, que comprende introducir una molécula de ácido nucleico de la invención en una planta transgénica, donde la molécula de ácido nucleico puede expresarse en la planta transgénica en una cantidad eficaz para el control de insectos. En una realización preferida, los

insectos son insectos coleópteros, preferiblemente gusano de la raíz del maíz del oeste y gusano de la raíz del maíz del norte.

En otra realización preferida, la invención abarca un método para controlar insectos, que comprende liberar a los insectos una cantidad eficaz de una toxina Cry3A modificada de la invención. De acuerdo con esta realización, los insectos son insectos coleópteros, preferiblemente gusanos de la raíz del maíz del oeste y gusanos de la raíz del maíz del norte. Preferiblemente la toxina Cry3A modificada es liberada a los insectos por vía oral. En un aspecto preferido la toxina es liberada oralmente a través de una planta transgénica que comprende una secuencia de ácido nucleico que expresa una toxina Cry3A modificada de la presente invención.

La presente invención abarca también un método para la preparación de una toxina Cry3A modificada que comprende: (a) obtener un gen de una toxina *cry*3A, que codifica una toxina Cry3A; (b) identificar una proteasa intestinal de un insecto objetivo, (c) obtener una secuencia nucleotídica que codifica un sitio de reconocimiento para la proteasa intestinal; (d) insertar la secuencia nucleotídica de (c) dentro del dominio I o del dominio III, o en ambos, dominios el I y el III, en una posición que reemplaza, que está dentro, o que es adyacente a la secuencia nucleotídica que codifica para un sitio de reconocimiento de proteasa natural en el gen de la toxina *cry*3A, creando de esta manera un gen de una toxina *cry*3A, modificada; (e) insertar el gen de la toxina *cry*3A modificada en un cassette de expresión; (f) expresar el gen de la toxina *cry*3A modificada en una célula huésped no humana, lo cual da como resultado una célula huésped que produce una toxina Cry3A modificada; y (g) bioensayar la toxina Cry3A modificada contra un insecto objetivo, lo cual produce una mortalidad más elevada para el insecto objetivo que la mortalidad causada por una toxina Cry3A. En una realización preferida, la toxina Cry3A modificada causa por lo menos aproximadamente 50%

de mortalidad para un insecto objetivo, mientras que la toxina Cry3A causa hasta aproximadamente 30% de mortalidad.

La presente invención abarca además un método para controlar insectos en el cual la planta transgénica comprende además una segunda secuencia de ácido nucleico o grupos de secuencias de ácido nucleico que codifican un segundo principio pesticida. Son segundas secuencias de ácido nucleico particularmente preferidas, aquellas que codifican una endotoxina δ, aquellas que codifican una toxina de Proteína Insecticida Vegetativa, descritas en las Patentes Norteamericanas N: 5.849.870 y 5.877.012, que se incorporan aquí como referencia, o aquellas que codifican una vía para la producción de un principio no proteináceo. En otras realizaciones las secuencias nucleotídicas de la invención pueden modificarse adicionalmente mediante la incorporación de mutaciones al azar en una técnica que se conoce como recombinación *in vitro* o entremezclado de ADN. Esta técnica ha sido descrita en Stemmer *et al.*, Nature 370:389-391 (1994) y en la Patente Norteamericana Nº 5.609.793, que se incorpora aquí como referencia. Se producen millones de copias mutantes de una secuencia nucleotídica en base a una secuencia nucleotídica original de esta invención y se recuperan variantes con propiedades mejoradas tales como con actividad insecticida incrementada, con estabilidad mejorada o con diferentes especificidades o rangos de plagas de insectos objetivo. El método abarca la formación de un polinucleótido mutagenizado de doble hélice a partir de un patrón de polinucleótido de doble hélice que comprende una secuencia nucleotídica de esta invención, donde el patrón de polinucleótido de doble hélice ha sido disociado en fragmentos al azar de doble hélice de un tamaño deseado, y que comprende las etapas de agregar a la población resultante de fragmentos al azar de doble hélice uno o más oligonucleótidos de doble hélice o de un solo filamento donde dichos oligonucleótidos comprenden un área de identidad y un área de

heterología para el polinucleótido de patrón de doble hélice; desnaturalizar la mezcla resultante de los fragmentos al azar de doble hélice y los oligonucleótidos en fragmentos de un solo filamento; incubar la población resultante de fragmentos de un solo filamento con una polimerasa bajo condiciones que dan como resultado el acoplamiento de dichos fragmentos de un solo filamento en dichas áreas de identidad para formar pares de fragmentos acoplados, siendo dichas áreas de identidad suficientes para que uno de los miembros de un par cebe la replicación del otro, formando de este modo un polinucleótido de doble filamento mutagenizado; y repetir la segunda y tercera etapas durante por lo menos dos ciclos más, donde la mezcla resultante en la segunda etapa de un ciclo adicional incluye el polinucleótido mutagenizado de doble filamento a partir de la tercera etapa del ciclo previo, y el ciclo adicional forma otro polinucleótido de doble filamento mutagenizado. En una realización preferida, la concentración de una sola especie de fragmento al azar de doble filamento en la población de fragmentos al azar de doble filamento es inferior al 1% en peso del ADN total. En otra realización preferida, el patrón de polinucleótido de doble hélice comprende por lo menos aproximadamente 100 especies de polinucleótidos. En otra realización preferida, el tamaño de los fragmentos al azar de doble hélice es de aproximadamente 5 pb hasta 5 kb. En otra realización preferida, la cuarta etapa del método comprende repetir la segunda y tercera etapa durante por lo menos 10 ciclos.

Expresión de la Secuencia Nucleotídica en Huéspedes Microbianos Heterólogos

Como agentes biológicos para el control de insectos, las toxinas Cry3A insecticidas modificadas, se producen mediante la expresión de las secuencias nucleotídicas en células huésped heterólogas que son capaces de expresar las secuencias nucleotídicas. En una primera realización, se preparan células de *B.*

thuringiensis que comprenden modificaciones de una secuencia nucleotídica de esta invención. Dichas modificaciones abarcan mutaciones o deleciones del elementos reguladores existentes, lo cual conduce a la expresión alterada de la secuencia nucleotídica, o a la incorporación de nuevos elementos reguladores que controlan la expresión de la secuencia nucleotídica. En otra realización, se agregan copias adicionales de una o más secuencias nucleotídicas a las células de *B. thuringiensis* mediante la inserción en el cromosoma o mediante la introducción de moléculas que se replican extracromosomalmente que contienen las secuencias nucleotídicas.

En otra realización, por lo menos una de las secuencias nucleotídicas de la invención se inserta dentro de un cassette de expresión apropiado, que comprende una señal promotora y de terminación. La expresión de la secuencia nucleotídica es constitutiva, o se usa un promotor inducible que responde a varios tipos de estímulos para iniciar la transcripción. En una realización preferida, la célula en la cual se expresa la toxina es un microorganismo, tal como un virus, una bacteria o un hongo. En una realización preferida, un virus, tal como un baculovirus contiene una secuencia nucleotídica de la invención en su genoma y expresa grandes cantidades de la correspondiente toxina insecticida después de la infección de las células eucarióticas que son apropiadas para la replicación del virus y para la expresión de la secuencia nucleotídica. La toxina insecticida así producida se usa como agente insecticida. Alternativamente, los baculovirus manipulados por ingeniería genética para que incluyan la secuencia nucleotídica, se usan para infectar insectos *in vivo* y para matarlos, ya sea mediante la expresión de la toxina insecticida o mediante una combinación de una infección viral y la expresión de la toxina insecticida.

La células bacterianas son también huéspedes para la expresión de la secuencia nucleotídica de la invención. En una realización preferida, se usan

bacterias simbióticas no patógenas, que son capaces de vivir y replicarse entre tejidos de plantas, denominadas endofitos, o bacterias simbióticas no patógenas, que son capaces de colonizar la filoesfera o la rizoesfera, denominados epifitos. Dichas bacterias incluyen bacterias de los géneros *Agrobacterium, Alcaligenes, Azospirillum, Azotobacter, Bacillus, Clavibacter, Enterobacter, Erwinia, Flavobacter, Klebsiella, Pseudomonas, Rhizobium, Serratia, Streptomyces y Xanthomonas.* También son huéspedes posibles para la expresión de las secuencias nucleotídicas de la invención para el mismo propósito, hongos simbióticos tales como *Trichoderma y Gliocladium.* Las técnicas para estas manipulaciones genéticas son específicas para los diferentes huéspedes disponibles y son conocidas en el arte. Por ejemplo, los vectores de expresión pKK223-3 y pKK223-2 pueden usarse para expresar los genes heterólogos en *E. coli* ya sea por fusión de transcripción o de traducción atrás del promotor tac o trc. Para la expresión de los operones que codifican los ORFs múltiples, el procedimiento más simple consiste en insertar el operón dentro de un vector tal como pKK223-3 en una fusión de transcripción, lo cual permite usar el sitio de unión del ribosoma relacionado de los genes heterólogos. Las técnicas para la hiperexpresión en las especies gram-positivas, tales como *Bacillus* son también conocidas en el arte y pueden usarse en el contexto de esta invención (Quax et al. In:Industrial Microorganisms:Basic and Applied Molecular Genetics, Eds. Baltz *et al.*, American Society por Microbiology, Washington (1993)). Los sistemas alternativos para la hiperexpresión se basan por ejemplo en los vectores de levadura e incluyen el uso de *Pichia, Saccharomyces* y *Kluyveromyces* (Sreekrishna, In:Industrial microorganisms:basic and applied molecular genetics, Baltz, Hegeman, and Skatrud eds., American Society for Microbiology, Washington (1993); Dequin & Barre, Biotechnology L2:173-177 (1994); van der Berg *et al.*, Biotechnology 8:135-139 (1990)).

Transformación de plantas

En una realización particularmente preferida, por lo menos una de las toxinas Cry3A modificadas insecticidas de la invención se expresa en un organismo superior por ejemplo, en una planta. En este caso, las plantas transgénicas que expresan cantidades eficaces de las toxinas Cry3A modificadas se protegen a sí mismas de las plagas de insectos. Cuando el insecto comienza a comer dicha planta transgénica, ingiere también las toxinas Cry3A modificadas expresadas. Esto impide que el insecto siga mordiendo el tejido de la planta o incluso que pueda dañar o matar al insecto. Una secuencia nucleotídica de la presente invención se inserta dentro de un cassette de expresión, el cual es luego integrado preferiblemente en forma estable en el genoma de dicha planta. En otra realización preferida, la secuencia nucleotídica se incluye en un virus auto-replicante no patógeno. Las plantas transformadas de acuerdo con la presente invención pueden ser monocotiledóneas o dicotiledóneas e incluyen pero no están limitadas a, maíz, trigo, cebada, centeno, batata, frijoles, guisantes, achicoria, lechuga, repollo, coliflor, brócoli, rábano, espinaca, espárrago, cebolla, ajo, pimiento, perejil, calabaza, zapallo, cáñamo, zucchini, manzana, pera, membrillo, melón, ciruelo, cerezo, melocotón, mandarina, albaricoque, frutilla, uva, frambuesa, mora, piña, palta, papaya, mango, banana, soja, tomate, sorgo, caña de azúcar, remolacha, girasol, colza, trébol, tabaco, zanahoria, algodón, alfalfa, arroz, patata, berenjena, pepino, Arabidopsis, y plantas leñosas tales como coníferas y árboles de hojas caducas. Una vez transformada la secuencia nucleotídica deseada en una especie de planta particular, dicha planta puede propagarse en esa especie o pasarse a otras variedades de la misma especie, incluyendo en particular variedades comerciales, usando técnicas de cultivo tradicionales. Una secuencia nucleotídica de esta invención se expresa preferiblemente en plantas transgénicas, causando de esta manera la biosíntesis

de la correspondiente toxina Cry3A modificada en las plantas transgénicas. De esta manera, se generan plantas transgénicas con resistencia mejorada a los insectos. Para su expresión en plantas transgénicas, las secuencias nucleotídicas de la invención pueden requerir otras modificaciones y la optimización. Aunque en muchos casos los genes de organismos microbianos pueden expresarse en plantas a niveles elevados sin modificación, la baja expresión en las plantas transgénicas, puede ser el resultado de secuencias nucleotídicas microbianas que tienen codones que no son preferidos en las plantas. Es conocido en el arte que todos los organismos tienen preferencia específica para el uso de codones, y los codones de las secuencias nucleotídicas descritos en esta invención pueden cambiarse para adaptarse a las preferencias de la planta, manteniendo al mismo tiempo los aminoácidos codificados por los mismos. Además, se logra una mejor expresión elevada en las plantas, a partir de las secuencias codificadoras que tienen por lo menos aproximadamente un contenido de 35% de GC, preferiblemente más de aproximadamente 45%, más preferiblemente más de aproximadamente el 50% y aún más preferiblemente más de aproximadamente el 60%. Las secuencias nucleotídicas microbianas que tienen bajos contenidos de GC pueden expresarse muy pobremente en las plantas debido a la existencia de motivos ATTTA que pueden desestabilizar los mensajes, y motivos AATAAA que pueden causar una poliadenilación inapropiada. Aunque las secuencias génicas preferidas pueden expresarse adecuadamente, tanto en las especies de plantas monocotiledóneas como en las dicotiledóneas, las secuencias pueden modificarse para tener en cuenta las preferencias de codones específicos y las preferencias de contenido de GC de monocotiledones o dicotiledones ya que se ha demostrado que estas preferencias difieren (Murray *et al.*, Nucl. Acids. Res. 17:477-498 (1989)). Además, las secuencias nucleotídicas son rastreadas para determinar la existencia de sitios de empalme ilegítimos que pueden causar el

truncamiento del mensaje. Todos los cambios que son necesarios dentro de las secuencias nucleotídicas tales como los descritos anteriormente, se efectúan usando técnicas bien conocidas de mutagénesis dirigida al sitio, PCR, y construcción génica sintética usando los métodos descritos en las solicitudes de patentes publicadas EP 0 385 962 (to Monsanto), EP 0 359 472 (to Lubrizol, y WO 93/07278 (de Ciba-Geigy).

En una realización de la invención, se prepara un gen *cry*3A de acuerdo con el procedimiento descrito en la Patente Norteamericana Nº 5.625.136, que se incorpora aquí como referencia. En este procedimiento se usan codones preferidos de maíz, es decir, un codón único que muy frecuentemente codifica el aminoácido en el maíz. El codón preferido del maíz para un aminoácido particular puede derivar por ejemplo de secuencias génicas conocidas del maíz. El uso del codón del maíz para 28 genes de plantas de maíz se encuentra en Murray *et al.*, Nucleic Acids Research 17:477- 498 (1989), cuya descripción se incorpora aquí como referencia. Una secuencia de síntesis preparada con codones optimizados de maíz se establece en la SEC ID NO: 3.

De esta manera, las secuencias nucleotídicas pueden optimizarse para la expresión en cualquier planta. Se ha reconocido que todas o cualquier parte de la secuencia génica puede ser optimizada o sintetizada. Es decir, pueden usarse también secuencias sintéticas o parcialmente optimizadas.

Para una iniciación de traducción eficiente, las secuencias adyacentes a la metionina iniciadora pueden requerir modificación. Por ejemplo, pueden modificarse mediante la inclusión de secuencias que se sabe que son eficaces en plantas. Joshi ha sugerido un concenso apropiado para plantas (NAR 15:6643-6653 (1987)) y Clonetech sugiere un iniciador más de traducción de consenso (1993/1994 catálogo, página 210). Estos consensos son apropiados para ser usados con las secuencias nucleotídicas de esta invención. Las secuencias se

incorporan en las construcciones que comprenden las secuencias nucleotídicas, hasta e incluyendo el ATG (al mismo tiempo que deja al segundo aminoácido sin modificar), o alternativamente hasta e incluyendo el GTC subsiguiente al ATG (con la posibilidad de modificar el segundo aminoácido del transgen).

La expresión de la secuencias nucleotídicas en plantas transgénicas, es dirigida por los promotores que funcionan en las plantas. La elección del promotor variará dependiendo de los requisitos temporales y espaciales para la expresión, y dependiendo también de la especie objetivo. Por lo tanto, se prefiere la expresión de las secuencias nucleotídicas de esta invención en las hojas, en los tallos, o en los vástagos, en las mazorcas, en las inflorescencias (por ejemplo espigas, panículos, mazorcas de maíz, etc), en las raíces, y/o en las plántulas. En muchos casos, sin embargo, se busca protección contra más de un tipo de plaga de insectos y por lo tanto es deseable la expresión en múltiples tejidos. Aunque muchos promotores para dicotiledones han demostrado que son operacionales en los monocotiledones y viceversa, idealmente los promotores dicotiledóneos se seleccionan para la expresión en los dicotiledones, y los promotores monocotiledóneos para la expresión en los monocotiledones. Sin embargo, no hay ninguna restricción en cuando a la procedencia de los promotores seleccionados; es suficiente que sean operacionales para dirigir la expresión de las secuencias nucleotídicas en la célula deseada.

Los promotores preferidos que se expresan constitutivamente incluyen promotores de genes que codifican actina o ubiquitina y los promotores caMV 35S y 19S. La secuencia nucleotídica de esta invención puede expresarse también bajo la regulación de promotores que están regulados químicamente. Esto permite que las toxinas de Cry3A insecticidas modificadas puedan ser únicamente sintetizadas, cuando se tratan las plantas de cultivo con los productos químicos inductores. La tecnología preferida para la inducción quimica de la

expresión génica se detalla en la solicitud publicada EP 0 332 104 (de Ciba Geigy) y la Patente Norteamericana N 5.614.395. Un promotor preferido para la inducción química es el promotor PR-1a del tabaco.

Una categoría de promotores preferida es la que es inducible por heridas. Se han descrito numerosos promotores que se expresan en sitios con heridas y también en sitios de infección con fitopatógenos. Idealmente, dicho promotor debería ser únicamente activo localmente en los sitios de infección, y de esta manera las toxinas insecticidas Cry3A modificadas se acumularían únicamente en células que necesitan sintetizar las toxinas Cry3A insecticidas modificadas para matar a la plaga de insectos invasores. Los promotores de esta clase preferidos incluyen los descritos por Stanford *et al.*, Mol. Gen. Genet. 215:200-208 (1989), Xu et al., Plant Molec. Biol. 22:573-588 (1993), Logemann *et al.* Plant Cell 1:151-158 (1989), Rohrmeier & Lehl, Plant. Molec. Biol. 22:783-792 (1993), Firek *et al.* Plant Molec. Biol. 22:129-142 81993), y Warner *et al.* Plant. J. 3:191-201 (1993). Los promotores específicos de tejidos o preferenciales de tejidos, que son útiles para la expresión de los genes de la toxina Cry3A modificada en plantas, particularmente en el maíz, son aquellos que dirigen la expresión en las raíces, médula, hoja o polen, particularmente en la raíz. Dichos promotores, los que son aislados a partir de PEPC o trpA, han sido descritos en la Patente Norteamericana N: 5.625.136, o MTL, descritos en la Patente Norteamericana Nº 5.466.785. Ambas patentes norteamericanas se incorporan aquí como referencia en su totalidad.

Otras realizaciones preferidas son plantas transgénicas que expresan las secuencias nucleotídicas en una manera inducible por herida o inducible por infección de patógenos.

Además de los promotores, están también disponibles una variedad de terminadores de transcripción, para ser usados en la construcción de genes

150
152

quiméricos, usando los genes de la toxina Cry3A modificados de la presente invención. Los terminadores de transcripción son responsables de la terminación de transcripción más allá del transgen y de su correcta poliadenilación. Entre los terminadores transcripcionales apropiados y aquellos que se sabe que funcionan en plantas, se incluyen el terminador CaMV 35S, el terminador tml, el terminador de nopalina sintasa, el terminador rbcS E9 de guisante y otros que son conocidos en el arte. Estos pueden usarse tanto en los monocotiledones como en los dicotiledones. Cualquier terminador disponible que se sabe que funciona en las plantas puede ser utilizado en el contexto de esta invención.

Pueden incorporarse numerosas otras secuencias en los cassettes de expresión descritos en esta invención. Estos incluyen secuencias que han demostrado que mejoran la expresión tal como secuencias de intrones (por ejemplo de Adhl y bronzel) y secuencias líder virales (por ejemplo, de TMV, MCMV y AMV).

Puede ser preferible dirigir la expresión de las secuencias nucleotídicas de la presente invención a diferentes localizaciones celulares en la planta. En algunos casos, la localización en el citosol puede ser deseable, mientras que en otros casos puede preferirse la localización en alguna organela sub-celular. La localización sub-celular de las enzimas codificadas por transgenes, se lleva a cabo usando técnicas bien conocidas en el arte. Típicamente el ADN que codifica al péptido objetivo de un producto génico conocido, que tiene como blanco a la organela, es manipulado y fusionado cadena arriba de la secuencia nucleotídica. Muchas de dichas secuencias objetivo son conocidas para el cloroplasto, y se ha mostrado su funcionamiento en construcciones heterólogas. La expresión de las secuencias nucleotídicas de la presente invención tiene también como blanco al retículo endoplásmico o a las vacuolas de las células huésped. Las técnicas para lograr esto son bien conocidas en el arte.

Los vectores que son apropiados para la transformación de plantas se describen en otra parte de esta memoria. Para las transformaciones intermediadas por Agrobacterium, son apropiados vectores binarios o vectores portadores de por lo menos una secuencia límite de T- ADN, mientras que para la transferencia génica directa cualquier vector es apropiado, y puede preferirse el ADN lineal que contiene únicamente la construcción de interés. En el caso de transferencia génica directa, puede usarse la transformación con una sola especie de ADN, o la co-transformación (Schocher et al., Biotechnology 4:1093-1096 (1986)). Para ambas, la transferencia génica directa y la transferencia intermediada por Agrobacterium, la transformación es usualmente (pero no necesariamente) efectuada con un marcador seleccionable que puede proveer resistencia a un antibiótico (kanamicina, higromicina, o metotrexato) o un herbicida (basta). Los vectores de transformación de plantas que comprenden los genes de la toxina Cry3A modificados de la presente invención, pueden comprender también genes (por ejemplo fosfomanosa isomerasa, PMI) que proporcionan selección positiva de las plantas transgénicas tal como se describió en las Patentes Norteamericanas 5.767.378 y 5.994.629, que se incorporan aquí como referencia. La elección del marcador seleccionable no es sin embargo crítico para la invención.

En otra realización, una secuencia nucleotídica de la presente invención se transforma directamente al genoma plástido. Una ventaja importante de la transformación del plástido es que los plástidos generalmente son capaces de expresar genes bacterianos sin optimización sustancial del codón, y que los plástidos son capaces de expresar múltiples marcos de lectura abierta bajo el control de un solo promotor. La tecnología de transformación de plástidos, ha sido extensamente descrita en las Patentes Norteamericanas Nº 5.451.513, 5.545.817 y 5.545.818, en la solicitud PCT no. WO 95/16783. La técnica básica para la

transformación de cloroplastos involucra introducir regiones del ADN plástido clonado que flanquea un marcador seleccionable conjuntamente con el gen de interés en un tejido objetivo apropiado, por ejemplo usando biolística o transformación de protoplastos (por ejemplo cloruro de calcio o transformación intermediada por PEG). Las regiones flanqueadoras de 1 a 1,5 kb, denominadas secuencias para hacer blanco en el objetivo, facilitan la recombinación homóloga con el genoma plástido y permiten por lo tanto el reemplazo o modificación de las regiones específicas del plastoma. Inicialmente, las mutaciones puntuales en el ARNr 16S y el cloroplasto, y los genes rps 12 que confieren resistencia a la espectinomicina y/o a la estreptomicina se utilizan como marcadores seleccionables para la transformación (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). Esto da como resultado transformantes homoplásmicos estables, a una frecuencia de aproximadamente uno por cada 100 bombardeos de hojas-objetivo. La presencia de los sitios de clonación entre estos marcadores, permite la creación de un vector plástido objetivo, para la introducción de genes extraños (Staub, J.M. and Maliga, P. (1993) EMBO J. 12, 601-606). Se obtienen aumentos sustanciales en la frecuencia de transformación mediante el reemplazo del ARNr recesivo o los genes de resistencia al antibiótico de proteína r, con un marcador dominante seleccionable, donde el gen aadA bacteriano codifica la enzima cletoxificadora de espectomicina aminoglicosido-3'-adeniltransferasa (Svab, Z., y Maliga, P. (1993) Proc. Natl. Acad. Sci. USA 90, 913-917). Previamente, este marcador había sido exitosamente usado para la transformación de alta frecuencia del genoma plástido de las algas verdes Chlamydomonas reinhardtii (Goldschmidt-Clermont, M. (1991) Nucl. Acids. Res. 19:4083-4089). Otros marcadores seleccionables útiles para la tranformación de plástidos son conocidos en el arte y están abarcados dentro del alcance de la

LEG
155

invención. Típicamente, se requieren aproximadamente 15-20 ciclos de división celular a continuación de la transformación para alcanzar un estado homoplastídico. La expresión del plástido, en el cual se insertan los genes mediante recombinación homológa dentro de todas las varias miles de copias del genoma plástido circular, presente en cada célula de planta, aprovecha la enorme ventaja de la cantidad de copias con respecto a los genes expresados en forma nuclear, para permitir niveles de expresión que exceden fácilmente el 10% de la proteína total de planta. En una realización preferida, una secuencia nucleotídica de la presente invención se inserta en un vector para el objetivo plástido y se transforma en el genoma plástido de un huésped de planta deseado. Se obtienen plantas homoplásticas para genomas plástidos que contienen una secuencia nucleotídica de la presente invención, y que son preferiblemente capaces de una elevada expresión de la secuencia nucleotídica.

Combinaciones de Principios para el Control de Insectos

Las toxinas Cry3A modificadas de la invención pueden usarse en combinación con las endotoxinas δ del Bt o con otros principios pesticidas, para aumentar la tasa que hace blanco en la plaga. Además, el uso de las toxinas Cry3A modificadas de la invención en combinación con las endotoxinas δ Bt o con otros principios pesticidas de distinta naturaleza, tiene una utilidad particular para la prevención y/o manipulación de la resistencia a los insectos.

Otros principios insecticidas incluyen por ejemplo lectinas, α-amilasa, peroxidasa y colesterol oxidasa. Los genes de Proteína Insecticida Vegetativa, tales como *vip*1A(a) y *vip*2A(a) descritos en la Patente Norteamericana N: 5.889.174 y que se incorporan aquí como referencia, son también útiles en la presente invención.

Esta co-expresión de más de un principio insecticida en la misma planta transgénica puede lograrse manipulando por ingeniería genética una planta de

manera que contengan y exprese todos los genes necesarios. Alternativamente, una planta, parental 1, puede manipularse genéticamente para que exprese los genes de la presente invención. Una segunda planta, parental 2, puede manipularse genéticamente para que exprese un principio de control de insectos suplementario. Cruzando el Parental 1 con el Parental 2, se obtienen plantas de una progenie que expresa todos los genes introducidos en las plantas Parental 1 y Parental 2. Las semillas transgénicas de la presente invención pueden tratarse también con un recubrimiento insecticida para las semillas tal como se describe en las Patentes Norteamericanas N: 5.849.320 y 5.876.739, que se incorporan aquí como referencia. Cuando tanto el recubrimiento insecticida para la semilla como la semilla transgénica de la invención son activas contra el mismo insecto objetivo, la combinación es útil (i) en un método para mejorar la actividad de una toxina Cry3A modificada de la invención contra el insecto objetivo y (ii) en un método para evitar el desarrollo de la resistencia a una toxina Cry3A modificada de la invención proporcionando un segundo mecanismo de acción contra el insecto objetivo. Por lo tanto, la invención provee un método para mejorar la actividad contra o para prevenir el desarrollo de la resistencia en un insecto objetivo, por ejemplo en un gusano de la raíz del maíz, que comprende aplicar un recubrimiento de una semilla insecticida a una semilla transgénica que comprende una o más toxinas Cry3A modificadas de la invención.

Incluso cuando el recubrimiento insecticida de la semilla es activo contra un insecto diferente, el recubrimiento insecticida de la semilla es útil para expandir la tasa de control de insectos, por ejemplo, mediante el agregado de un recubrimiento insecticida para la semilla, que tiene actividad contra los insectos lepidópteros, a la semilla transgénica de la invención, que tiene actividad contra insectos coleópteros, la semilla transgénica recubierta produjo controles tanto para plagas de insectos lepidópteros como coleópteros.

EJEMPLOS

La invención se describirá adicionalmente con referencia a los siguientes ejemplos detallados. Estos ejemplos se proveen con propósitos ilustrativos únicamente y no están destinados a ser limitativos a menos que se especifique lo contrario. Las técnicas de clonación molecular y de ADN recombinante convencionales que se usaron son bien conocidas en el arte y han sido descritas por J. Sambrook, et al., Molecular Cloning; *A. Laboratory Manual, 3d Ed.*, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press (2001); by T.J. Silhavy, M.L. Berman, and L.W. Enquist, *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) and by Ausubel, F.M. et al., *Current Protocols in Molecular Biology*, New Cork, John Wiley and Sons Inc., (1988), Reiter, et al., *Methods in Arabidopsis Research*, World Scientific Press (1992), and Schultz et al., *Plant Molecular Biology Manual*, Kluwer Academic Publishers (1998).

Ejemplo 1: **Construcción Génica *cry*3A Optimizado de Maíz**

El gen *cry*3A optimizado de maíz se preparó de acuerdo con el procedimiento descrito en la Patente Norteamericana N. 5.625.136. En este procedimiento, se usan codones preferidos en el maíz, es decir, el codón único que muy frecuentemente codifica el aminoácido del maíz. El codón preferido del maíz para un aminoácido particular deriva de secuencias génicas conocidas del maíz. El uso del codón del maíz para 28 genes de plantas de maíz se encuentra en Murray *et al.*, Nucleic Acids Research 17:477-498 (1989). La secuencia sintética preparada con codones optimizados de maíz se indica en la SEC ID NO: 3.

Ejemplo 2: **Identificación de la Actividad Enzimática de la Catepsina G en Intestinos del Gusano de la Raíz del Maíz del Oeste.**

Las actividades enzimáticas de la Catepsina de tipo G (serina proteasa) y del tipo de la Catepsina B (cisteína proteasa) en los intestinos del gusano de la raíz del maíz del oeste se midieron usando sustratos colorimétricos. Cada reacción de 1 ml contiene cinco mesenterones homogeneizados del tercer estadio del gusano de la raíz del maíz del oeste, y 1 mg de sustrato disuelto en buffer de reacción (10 mM de Tris, 5 mM de NaCl, 0,01 M de DTT, pH 7,5). El sustrato de catepsina G ensayado es Ala-Ala-Pro-Phe (SEC ID NO: 35)-*p*NA y sustrato de catepsina B, Arg-Arg-*p*NA. Las reacciones se incubaron a 28°C durante una hora. La intensidad de la formación de color amarillo, indicativa de la eficiencia de una proteasa para reconocer el sustrato apropiado, se comparó en tratamientos versus controles. Las reacciones se evaluaron como negativas (-) si no se detectaba ningún color o si se detectaba solamente un ligero color de base. La reacciones que estaban un 25%, 50%, 75% o 100% por arriba de la base, se evaluaron como +, ++, +++, o ++++, respectivamente.

Los resultados de los ensayos enzimáticos se muestran en la Tabla siguiente:

Tabla 1

Reacción	Intensidad de Color del Producto
Intestino de WCR solamente	-
Sustrato de catepsina B solamente	-
Sustrato de catepsina G solamente	-
Sustrato de intestino de WCR + catepsina B	+
Sustrato de intestino de WCR + catepsina G	+++

Esta es la primera vez que se ha identificado la actividad de la catepsina G de serina proteasa en los intestinos de los gusanos de la raíz del maíz del oeste. Los intestinos de los gusanos de la raíz del maíz del oeste, tienen claramente una actividad más fuerte de la catepsina G, la serina proteasa, en comparación con la actividad de la catepsina B, la cisteína proteasa. La secuencia AAPF se selecciona como el sitio de reconocimiento de la proteasa de la catepsina G para crear las toxinas Cry3A modificadas de la presente invención.

Ejemplo 3: **Construcción de los Genes *cry*3A Modificados**

Los genes *cry*3A modificados que comprenden una secuencia nucleotídica que codifica el sitio de reconocimiento de la catepsina G en el dominio I, el dominio III y los dominios I y III, se prepararon usando PCR superpuesto.El gen *cry*3A optimizado de maíz (SEC ID NO: 2), comprendido en el plásmido pCIB6850 (SEC ID NO: 5), se usó como modelo de partida. Se prepararon ocho construcciones del gen *cry*3A modificados que codifican las toxinas Cry3A modificadas; *cry*3A054, *cry*3A055 y *cry*3A085, que comprenden la secuencia codificadora del sitio de reconocimiento de la catepsina G en el dominio I; *cry*3A058, *cry*3A082, que comprenden la secuencia codificadora del sitio de reconocimiento de la catepsina G en el dominio III; *cry*3A056, *cry*3A057 y *cry*3A083, que comprenden la secuencia codificadora del sitio de reconocimiento de la catepsina G en el dominio I y en el dominio III. Los ocho genes *cry*3A modificados, y las toxinas Cry3A modificadas a las cuales codifican se describen a continuación:

<u>*cry*3A054 comprendido en pCMS054</u>

El *cry*3A054 (SEC ID NO: 6) comprende una secuencia nucleotídica que codifica una toxina Cry3A modificada. Se usan tres pares cebadores de PCR

superpuestos para insertar la secuencia nucleotídica que codifica el sitio de reconocimiento de la catepsina G en el *cry*3A optimizado de maíz no modificado:

1. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3' (SEC ID NO: 22)
AAPFtail3 – 5'-GAACGGTGCAGCGGGGTTCTTCTGCCAGC-3' (SEC ID NO: 23)

2. Tail5mod – 5'-GCTGCACCGTTCCCCCACAGCCAGGGCCG-3' (SEC ID NO: 24)
XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3' (SEC ID NO: 25)

3. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3' (SEC ID NO: 22)
XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3' (SEC ID NO: 25)

El par cebador 1 y el par cebador 2 generan dos productos de PCR únicos. Estos productos se combinan luego en partes iguales, y se usa el par cebador 3 para unir los productos con el fin de generar un fragmento de PCR que es nuevamente clonado en el patrón pCIB6850 original. El gen *cry*3A054 modificado es luego transferido a pBluescript (Stratagene). El plásmido resultante se designa pCMS054 y comprende el gen *cry*3A054 (SEC ID NO: 6).

La toxina Cry3A054 modificada (SEC ID NO: 7), codificada por el gen *cry*3A modificado comprendido en pCMS054, tiene un sitio de reconocimiento de la catepsina G, que comprende la secuencia de aminoácidos AAPF (SEC ID NO: 35), que está insertada en el dominio I entre los aminoácidos 107 y 113 de la toxina Cry3A (SEC ID NO: 4) no modificada. El sitio de reconocimiento de la catepsina G reemplaza el sitio de reconocimiento de la tripsina natural y es adyacente a un sitio de reconocimiento de la quimotripsina natural.

*cry*3A055 comprendido en pCMS055

El *cry*3A055 (SEC ID NO: 8) comprende una secuencia nucleotídica que codifica una toxina Cry3A modificada. Se usan tres pares cebadores de PCR

superpuestos para insertar la secuencia nucleotídica que codifica el sitio de reconocimiento de la catepsina G en el *cry*3A optimizado de maíz no modificado:

1. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3'	(SEC ID NO: 22)
AAPFtail3 – 5'-GAACGGTGCAGCGGGGTTCTTCTGCCAGC-3'	(SEC ID NO: 23)
2. AAPFtail4 – 5'-GCTGCACCGTTCCGCAACCCCCACAGCCA-3'	(SEC ID NO: 26)
XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3'	(SEC ID NO: 25)
3. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3'	(SEC ID NO: 22)
XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3'	(SEC ID NO: 25)

El par cebador 1 y el par cebador 2 generan dos productos de PCR únicos. Estos productos se combinan luego en partes iguales, y se usa el par cebador 3 para reunir los productos con el fin de generar un fragmento de PCR que es nuevamente clonado en el patrón pCIB6850 original. El gen *cry*3A055 modificado es luego transferido a pBluescript (Stratagene). El plásmido resultante se designa pCMS055 y comprende el gen *cry*3A055 (SEC ID NO: 8).

La toxina Cry3A055 modificada (SEC ID NO: 9), codificada por el gen *cry*3A modificado comprendido en pCMS055, tiene un sitio de reconocimiento de catepsina G, que comprende la secuencia de aminoácido AAPF (SEC ID NO: 35), que está insertado en el dominio I entre los aminoácidos 107 y 111 de la toxina Cry3A (SEC ID NO: 4) no modificada. El sitio de reconocimiento de la catepsina G es adyacente a un sitio de reconocimiento de la tripsina y la quimotripsina natural.

<u>*cry*3A058 comprendido en pCMS058</u>

El *cry*3A058 (SEC ID NO: 14) comprende una secuencia nucleotídica que codifica una toxina Cry3A modificada. Se usan tres pares cebadores de PCR

superpuestos para insertar la secuencia nucleotídica que codifica el sitio de reconocimiento de la catepsina G en el *cry*3A optimizado de maíz no modificado:

1. SalExt – 5'-GAGCGTCGACTTCTTCAAC-3' (SEC ID NO: 27)
AAPF-Y2 – 5'-GAACGGTGCAGCGTATTGGTTGAAGGGGGC-3' (SEC ID NO: 28)

2. AAPF-Y1 – 5'-GCTGCACCGTTCTACTTCGACAAGACCATC-3' (SEC ID NO: 29)
SacExt – 5'-GAGCTCAGATCTAGTTCACGG-3' (SEC ID NO: 30)

3. SalExt – 5'-GAGCGTCGACTTCTTCAAC-3' (SEC ID NO: 27)
SacExt – 5'-GAGCTCAGATCTAGTTCACGG-3' (SEC ID NO: 30)

El par cebador 1 y el par cebador 2 generan dos productos de PCR únicos. Estos productos se combinan luego en partes iguales, y se usa el par cebador 3 para reunir los productos con el fin de generar un fragmento de PCR que es nuevamente clonado en el patrón pCIB6850 original. El gen *cry*3A058 modificado es luego transferido a pBluescript (Stratagene). El plásmido resultante se designa pCMS058 y comprende el gen *cry*3A058 (SEC ID NO: 14).

La toxina Cry3A058 modificada (SEC ID NO: 15), codificada por el gen *cry*3A, tiene un sitio de reconocimiento de catepsina G, que comprende la secuencia de aminoácido AAPF (SEC ID NO: 35), que está insertado en el dominio III entre los aminoácidos 540 y 541 de la toxina Cry3A no modificada (SEC ID NO: 4). El sitio de reconocimiento de la catepsina G está dentro de un sitio de reconocimiento de la quimotripsina natural.

<u>pCMS082 que comprende *cry*3A082</u>

El *cry*3A082 (SEC ID NO: 12) comprende una secuencia nucleotídica que codifica una toxina Cry3A modificada. Se usó un par cebador de PCR de QuikChange Site Directed Mutagenesis para insertar la secuencia nucleotídica

que codifica el sitio de reconocimiento de la catepsina G en el *cry*3A optimizado de maíz.

BBmod1 – 5'-CGGGGCCCCCGCTGCACCGTTCTACTTCGACA-3' (SEC ID NO: 31)
BBmod2 – 5'-TGTCGAAGTAGAACGGTGCAGCGGGGGCCCCG-3' (SEC ID NO: 32)

El par cebador genera un producto de PCR único. Este producto se clonó nuevamente al patrón pCIB6850 original. El gen *cry*3A082 modificado es luego transferido a pBluescript (Stratagene). El plásmido resultante se designa pCMS082 y comprende el gen *cry*3A082 (SEC ID NO: 12).

La toxina Cry3A082 modificada (SEC ID NO: 13), codificada por el gen *cry*3A modificado, tiene un sitio de reconocimiento de catepsina G, que comprende la secuencia de aminoácido AAPF (SEC ID NO: 35), que está insertado en el dominio III entre los aminoácidos 539 y 542 de la toxina Cry3A no modificada (SEC ID NO: 4). El sitio de reconocimiento de la catepsina G reemplaza un sitio de reconocimiento de quimotripsina natural.

*cry*3A056 comprendido en pCMS056

El *cry*3A056 (SEC ID NO: 18) comprende una secuencia nucleotídica que codifica una toxina Cry3A modificada. Se usan seis pares de cebadores de PCR superpuestos para insertar dos sitios de reconocimiento de la catepsina G en el *cry*3A no modificado:

1. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3' (SEC ID NO: 22)
AAPFcola 3 – 5'-GAACGGTGCAGCGGGGTTCTTCTGCCAGC-3' (SEC ID NO: 23)

2. AAPFcola 4 – 5'-GCTGCACCGTTCCGCAACCCCCACAGCCA-3' (SEC ID NO: 26)
XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3' (SEC ID NO: 25)

3. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3' (SEC ID NO: 22)
XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3' (SEC ID NO: 25)

4. SalExt – 5'-GAGCGTCGACTTCTTCAAC-3' (SEC ID NO: 27)
AAPF-Y2 – 5'-GAACGGTGCAGCGTATTGGTTGAAGGGGGC-3' (SEC ID NO: 28)

5. AAPF-Y1 – 5'-GCTGCACCGTTCTACTTCGACAAGACCATC-3' (SEC ID NO: 29)
SacExt – 5'-GAGCTCAGATCTAGTTCACGG-3' (SEC ID NO: 30)

6. SalExt – 5'-GAGCGTCGACTTCTTCAAC-3' (SEC ID NO: 27)
SacExt – 5'-GAGCTCAGATCTAGTTCACGG-3' (SEC ID NO: 30)

El par cebador 1 y el par cebador 2 generan dos productos de PCR únicos. Estos productos se combinan luego en partes iguales, y se usa el par cebador 3 para reunir los productos con el fin de generar un fragmento de PCR que es nuevamente clonado al plásmido pCIB6850 original. El gen *cry*3A055 modificado es luego transferido a pBluescript (Stratagene). El plásmido resultante se designa pCMS055. El par cebador 4 y el par cebador 5 generan otro conjunto de fragmentos únicos, los cuales son unidos con otra PCR con el par cebador 6. Este fragmento se clona en el dominio III del gen *cry*3A055 modificado comprendido en pCMS055. El plásmido resultante se designa pCMS056 y comprende el gen *cry*3A056 (SEC ID NO: 18).

La toxina Cry3A056 modificada (SEC ID NO: 19), codificada por el gen *cry*3A modificado, tiene un sitio de reconocimiento de catepsina G, que

comprende la secuencia de aminoácido AAPF (SEC ID NO: 35), que está insertada en el dominio I entre los aminoácidos 107 y 111 y en el dominio III entre los aminoácidos 540 y 541 de la toxina Cry3A no modificada (SEC ID NO: 4). El sitio de reconocimiento de la catepsina G es adyacente a un sitio de reconocimiento de quimotripsina y tripsina natural en el dominio I y está dentro de un sitio de reconocimiento de quimotripsina natural en el dominio III.

*cry*3A057 comprendido en pCMS057

El *cry*3A057 (SEC ID NO: 16) comprende una secuencia nucleotídica que codifica una toxina Cry3A modificada. Se usan seis pares de cebadores de PCR superpuestos para insertar dos sitios de reconocimiento de la catepsina G en el *cry*3A no modificado:

1. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3' (SEC ID NO: 22)
AAPF$^{\text{Colin 3}}$ – 5'-GAACGGTGCAGCGGGGTTCTTCTGCCAGC-3' (SEC ID NO: 23)

2. Tail5mod – 5'-GCTGCACCGTTCCCCCACAGCCAGGGCCG-3' (SEC ID NO: 24)
XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3' (SEC ID NO: 25)

3. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3' (SEC ID NO: 22)
XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3' (SEC ID NO: 25)

4. SalExt – 5'-GAGCGTCGACTTCTTCAAC-3' (SEC ID NO: 27)
AAPF-Y2 – 5'-GAACGGTGCAGCGTATTGGTTGAAGGGGGC-3' (SEC ID NO: 28)

5. AAPF-Y1 – 5'-GCTGCACCGTTCTACTTCGACAAGACCATC-3' (SEC ID NO: 29)
SacExt – 5'-GAGCTCAGATCTAGTTCACGG-3' (SEC ID NO: 30)

6. SalExt – 5'-GAGCGTCGACTTCTTCAAC-3' (SEC ID NO: 27)
SacExt – 5'-GAGCTCAGATCTAGTTCACGG-3' (SEC ID NO: 30)

El par cebador 1 y el par cebador 2 generan dos productos de PCR únicos. Estos productos se combinan luego en partes iguales, y se usa el par cebador 3 para reunir los productos con el fin de generar un fragmento de PCR que es nuevamente clonado en el plásmido pCIB6850 original. El gen *cry*3A054 modificado es luego transferido a pBluescript (Stratagene). El plásmido resultante se designa pCMS054. El par cebador 4 y el par cebador 5 generan otro conjunto de fragmentos únicos, los cuales son reunidos con otra PCR con el par cebador 6. Este fragmento se clona en el dominio III del gen *cry*3A054 modificado comprendido en pCMS054. El plásmido resultante se designa pCMS057 y comprende el gen *cry*3A057 (SEC ID NO: 16).

La toxina Cry3A057 modificada (SEC ID NO: 17), codificada por el gen *cry*3A modificado, tiene un sitio de reconocimiento de catepsina G, que comprende la secuencia de aminoácido AAPF (SEC ID NO: 35), que está insertada en el dominio I entre los aminoácidos 107 y 113 y en el dominio III entre los aminoácidos 540 y 541 de la toxina Cry3A no modificada (SEC ID NO: 4). El sitio de reconocimiento de la catepsina G reemplaza un sitio de reconocimiento de tripsina natural y es adyacente a un sitio de reconocimiento de quimotripsina natural en el dominio I y está dentro de un sitio de reconocimiento de quimotripsina natural en el dominio III.

*cry*3A083 comprendido en pCMS083

El *cry*3A083 (SEC ID NO: 20) comprende una secuencia nucleotídica que codifica una toxina Cry3A modificada. Se usaron tres pares de cebadores de PCR superpuestos y un par cebador de PCR QuikChange Site Directed Mutagenesis para insertar dos sitios de reconocimiento de la catepsina G en el *cry*3A no modificado:

1. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3'	(SEC ID NO: 22)
AAPFcode3 – 5'-GAACGGTGCAGCGGGGTTCTTCTGCCAGC-3'	(SEC ID NO: 23)
2. AAPFcode4 – 5'-GCTGCACCGTTCCGCAACCCCCACAGCCA-3'	(SEC ID NO: 26)
XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3'	(SEC ID NO: 25)
3. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3'	(SEC ID NO: 22)
XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3'	(SEC ID NO: 25)
BBmod1 – 5'-CGGGGCCCCCGCTGCACCGTTCTACTTCGACA-3	(SEC ID NO: 31)
BBmod2 – 5'-TGTCGAAGTAGAACGGTGCAGCGGGGGCCCCG-3'	(SEC ID NO: 32)

El par cebador 1 y el par cebador 2 generan dos productos de PCR únicos. Estos productos se combinan luego en partes iguales, y se usa el par cebador 3 para reunir los productos con el fin de generar un fragmento de PCR que es nuevamente clonado en el plásmido pCIB6850 original. El gen *cry*3A055 modificado es luego transferido a pBluescript (Stratagene). El plásmido resultante se designa pCMS055. El par cebador 4 genera otro fragmento único que es clonado en el dominio III del gen *cry*3A modificado comprendido en pCMS055. El plásmido resultante se designa pCMS083 y comprende el gen *cry*3A083 (SEC ID NO: 20).

La toxina Cry3A083 modificada (SEC ID NO: 21), codificada por el gen *cry*3A modificado, tiene un sitio de reconocimiento de catepsina G, que comprende la secuencia de aminoácido AAPF (SEC ID NO: 35), que está insertada en el dominio I entre los aminoácidos 107 y 111 y entre los aminoácidos 539 y 542 de la toxina Cry3A no modificada (SEC ID NO: 4). El sitio de reconocimiento de la catepsina G es adyacente a una tripsina natural y a un sitio de reconocimiento de quimotripsina natural en el dominio I y reemplaza un sitio de reconocimiento de quimotripsina natural en el dominio III.

*Cry*3A085 comprendido en pCMS085

El gen *cry*3A085 (SEC ID NO: 10) comprende una secuencia codificadora de catepsina G en la misma posición que en el gen *cry*3A055 descrito anteriormente. El gen *cry*3A085 tiene 24 nucleótidos adicionales insertados en el extremo 5' que codifica los aminoácidos 41-47 de la secuencia de aminoácido deducida en la SEC ID NO: 2, así como también una metionina adicional. Los nucleótidos adicionales se insertan en el extremo 5' del gen *cry*3A055 usando el siguiente para cebador de PCR:

mo3Aext- 5'-GGATCCACCATGAACTACAAGGAGTTCCTCCGC-ATGACCGCCGACAAC-3' (SEC ID NO: 33)

CMS16 – 5'-CCTCCACCTGCTCCATGAAG-3' (SEC ID NO: 34)

La toxina Cry3A085 modificada (SEC ID NO: 11), codificada por el gen *cry*3A modificado, tiene un sitio de reconocimiento de catepsina G, que comprende la secuencia de aminoácido AAPF (SEC ID NO: 35), que está insertada en el dominio I entre los aminoácidos 107 y 111 de la toxina Cry3A modificada (SEC ID NO: 4) y tiene 8 residuos de aminoácidos adicionales en el N-término del cual el segundo residuo corresponde al aminoácido número 41 de la secuencia de aminoácido establecida en la (SEC ID NO: 2).

Ejemplo 4: **Actividad Insecticida de las Toxinas Cry3A Modificadas**

Las toxinas Cry3A modificadas se ensayaron para determinar la actividad insecticida contra el gusano de la raíz del maíz del oeste, el gusano de la raíz del norte y el gusano de la raíz del maíz del sur en bioensayos de insectos. Los bioensayos se llevaron a cabo usando un método de incorporación de dieta. Se cultivaron durante la noche clones de *E. coli* que expresan una de las toxinas de Cry3A modificadas de la invención. Se sonicaron 500 µl de un cultivo nocturno y luego se mezclaron con 500 µl de una dieta artificial fundida (Marrone et al (1985)

J. of Economic Entomology 78:290-293). Una vez solidificada la dieta, se pasa a una placa petri y se colocan sobre la dieta 20 gusanos de la raíz del maíz recientemente nacidos. Las placas petri se mantienen a 30°C. La mortalidad se registra después de 6 días. Todas las toxinas Cry3A modificadas causan 50-100% de mortalidad en el gusano de la raíz del maíz del norte y del oeste mientras que la toxina Cry3A no modificada causa 0-30% de mortalidad. Ninguna de las toxinas Cry3A modificadas tiene actividad contra el gusano de la raíz del maíz del sur.

Ejemplo 5: **Creación de Plantas de Maíz Transgénicas que Comprenden las Secuencias Codificadoras *cry*3A Modificadas.**

Se eligen tres genes cry3A modificados, el *cry*3A055, representativo de una modificación del dominio I, el *cry*3A058, representativo de una modificación de un dominio III, y el *cry*3A056, representativo de una modificación de un dominio I y dominio III, para la transformación de plantas de maíz. Un cassette de expresión que comprende una secuencia codificadora *cry*3A modificada es transferido a un vector apropiado para la transformación de maíz intermediado por Agrobacterium. Para este ejemplo, un cassette de expresión comprende además del gen *cry*3A modificado, el promotor MTL (Patente Norteamericana Nº 5.466.785) y el terminador nos que es conocido en el arte.

La transformación de embriones de maíz inmaduros se lleva a cabo esencialmente tal como ha sido descrito en Negrotto *et al.*, 2000, Plant Cell Reports 19: 798-803. Para este ejemplo, todos los constituyentes del medio son tal como se han descrito en Negrotto *et al., supra.* Sin embargo, pueden emplearse varios constituyentes del medio conocidos en el arte,

Los genes usados para la transformación son clonados en un vector apropiado para la tranformación del maíz. Los vectores usados en este ejemplo contienen el gen de fosfomanosa isomerasa (PMI) para la selección de líneas transgénicas (Negrotto *et al.* (2000) Plant Cell Reports 19:798-803).

El *Agrobacterium* cepa LBA4404 (pSB1) que contiene el plásmido de transformación de plantas es cultivado en un medio sólido YEP (extracto de levadura (5 g/L), peptona (10 g/L), NaCl (5g/L), 15g/l agar, pH 6,8) durante 2-4 días a 28°C. Aproximadamente 0,8X , se suspenden 10^9 de *Agrobacterium* en un medio LS-inf suplementado con 100 µM As (Negrotto *et al.*, (2000) Plant Cell Rep 19: 798-803). En este medio las bacterias son preinducidas durante 30-60 minutos.

Se extraen embriones inmaduros A188 u otro genotipo apropiado, de mazorcas de 8 – 12 días en LS-inf + 100 µM de As. Los embriones se enjuagan una vez con un medio de infección fresco. Luego se agrega una solución de *Agrobacterium* y los embriones se someten a agitación vortical durante 30 segundos y se dejan sedimentar con las bacterias durante 5 minutos. Los embriones son transferidos luego con el escutelo hacia arriba a un medio de LSAs y se cultivn en la oscuridad durante dos a tres días. Subsiguientemente, entre 20 y 25 embriones por placa petri son transferidos a un medio de LSDc suplementado con cefotaxina (250 mg/l) y con nitrato de plata (1,6 mg/l) y se cultivan en la oscuridad a 28°C durante 10 días. Los embriones inmaduros, que producen los callos embriogénicos son transferidos a un medio de LSD1M0,5S. Los cultivos se seleccionan en este medio durante 6 semanas con una etapa de sub-cultivo a las 3 semanas.

Los callos sobrevivientes son transferidos a un medio Reg1 suplementado con mannosa. Después del cultivo bajo la luz (régimen de 16 horas de luz/ 8 horas de oscuridad), los tejidos verdes son transferidos a un medio Reg2 sin reguladores de crecimiento y se incuban durante 1-2 semanas. Las plántulas son transferidas a cajas Magenta GA-7 (Magenta Corp, Chicago III) que contienen un medio Reg3 y se cultivan bajo la luz. Después de 2-3 semanas, las plantas se someten a pruebas para determinar la presencia de los genes de PMI, y los genes

*cry*3A modificados por PCR. Las plantas positivas del ensayo de PCR son transferidas al invernadero y se prueba su resistencia al gusano de la raíz del maíz.

Ejemplo 6: **Análisis de Plantas de Maíz Transgénico**

Eficacia del Gusano de la Raíz del Maíz

Bioensayo de Extracción de la Raíz

Se toman muestras de las plantas al ser transplantadas desde las cajas Magenta GA-7 al suelo. Esto permite tomar muestras de las raíces en un ambiente razonablemente estéril en relación a las condiciones del suelo. La toma de muestras consiste en cortar un pequeño trozo de raíz (de casi 2-4 cm de longitud), y colocarlo sobre phytagar enriquecido (12 g de phytagar, 9 g de sucrosa, 3 ml de sales de MS, 3 ml de vitaminas de MS, 3 ml de Nistatina (25 mg/ml), 7 ml de Cefotaxima (50 mg/ml), 7 ml de Aureomicina (50 mg/ml), 7 ml de Estreptomicina (50 mg/ml), y 7 ml de dH_2O, 600 ml) en una pequeña placa petri. Los controles negativos son plantas o bien transgénicas que son negativas a la PCR para el gen *cry*3A modificado del mismo experimento, o bien plantas no transgénicas (de un tamaño similar a las plantas de ensayo), que se cultivan en el phytotron. Si se toman muestras de las raíces de control del suelo, las muestras de la raíz se lavan con agua para eliminar los residuos de tierra, se sumergen en una solución de Nistatina (5 mg/ml), se sacan del baño, se secan con toallitas de papel y se colocan en una placa de phytagar.

Las muestras de raíces se inoculan con 10 larvas del primer estadio, sobre la superficie interna de la tapa de cada placa de phytagar y luego las tapas se sellan otra vez herméticamente. Las larvas se manipulan usando un pincel de punta fina. Después de inocular todas las placas, se coloca la bandeja de placas en la oscuridad a temperatura ambiente hasta la recolección de datos.

A los 3-4 días de la inoculación, se recogen los datos. Se calcula el porcentaje de mortalidad de las larvas junto con una evaluación visual del deterioro de la raíz. El deterioro por alimentación se evalúa como elevado, moderado, bajo o ausente y se le da un valor numérico de 3, 2, 1 o 0, respectivamente. Las muestras de la raíz que causaban por lo menos 40% de mortalidad y que tenían una evaluación de deterioro de 2 o menos se consideraron positivas.

Los resultados en la Tabla siguiente muestran que las plantas que expresan una toxina Cry3A modificada causan desde 40-100% de mortalidad a los gusanos de la raíz del maíz del oeste, mientras que las plantas de control causan 0-30% de mortalidad. Asimismo, las plantas que expresan una toxina Cry3A modificada sostienen un deterioro por alimentación significativamente inferior que las plantas de control.

Tabla 2

Evento T0	Toxinas Cry3A Modificada, Expresada	Porcentaje de Mortalidad por Planta A B C D E	Evaluación del Deterioro Promedio por Evento
240A7	Cry3A055	80 40 80 60	0,8
240B2	Cry3A055	60 60 60 80	1,25
240B9	Cry3A055	40 60 60 100	1
240B10	Cry3A055	80 40 60 60	1
240A15	Cry3A055	80 60 50 70 70	0,6
240A5	Cry3A055	60 80 60	0,33
240A9	Cry3A055	50 60 60 70 70	1,6
244A4	Cry3A058	50	1
244A7	Cry3A058	40 40 60	1,3

244A5	Cry3A058	50	1
244B7	Cry3A058	90	1
244B6	Cry3A058	50 40 60	1
243A3	Cry3A056	50 90 80 60	1,25
243A4	Cry3A056	50 80 60	1,7
243B1	Cry3A056	80 90	0,5
243B4	Cry3A056	70 60 50 80	1,5
245B2	Cry3A056	90 50 70 60	1
WT1	-	0 10 20 10 0	2,6
WT2	-	0 30 0 0 20	2,8

Biensayo de Planta Entera

Algunas plantas positivas identificadas usando un bioensayo de extracción de raíz descripto anteriormente se evalúan para determinar la resistencia del gusano de la raíz del maíz del oeste usando un bioensayo de planta entera. Las plantas se infestan generalmente dentro de los tres días después de completar el ensayo de extracción de raíz.

Se pre-incubaron huevos del gusano de la raíz del maíz del oeste de manera que la puesta de huevos ocurriera 2-3 días después de la inoculación de la planta. Los huevos se suspendieron en 0,2% de agar y se aplicaron al suelo alrededor de las plantas de ensayo aproximadamente a razón de 200 huevos/planta.

Dos semanas después de poner los huevos, las plantas fueron evaluadas para determinar el deterioro causado por las larvas del gusano de la raíz del maíz del oeste. Se usaron los criterios de altura de la planta, debilitamiento y masa de raíz para determinar si las plantas eran resistentes al deterioro por alimentación del gusano de la raíz del maíz del oeste. En el momento de la evaluación, las plantas de control eran típicamente más pequeñas que las plantas de Cry3A

modificadas. Asimismo, las plantas de control no transgénicas y las plantas que expresan las toxinas Cry3A no modificadas codificadas por el gen *cry*3A optimizado del maíz se debilitaron durante este tiempo debido a un severo cercenamiento de la mayor parte de las raíces, lo cual dio como resultado una falta de acumulación de masa en la raíz. En el momento de la evaluación, las plantas que expresan una toxina Cry3A modificada de la invención son más altas que las plantas de control, no están debilitadas, y tienen una masa de raíz intacta mayor debido a la actividad insecticida de la toxina Cry3A modificada.

Ensayo ELISA

Se usó el análisis ELISA de acuerdo con el método descrito en la Patente Norteamericana Nº 5.625.136 para la determinación cuantitativa del nivel de proteínas Cry3A modificada y no modificada en plantas transgénicas.

Tabla 3: Resultados del Bioensayo de Planta Entera y Niveles de Proteína

Planta de Maíz Transgénica	Tipo de Toxina Cry3A Expresada	Nivel de Proteínas Cry3A en las Raíces (ng/mg)	Planta Debilitada	Masa de Raíz Intacta
240A2E	Cry3A055 modificado	224	-	+
240A9C	Cry3A055 modificado	71	-	+
240B9D	Cry3A055 modificado	204	-	+
240B9E	Cry3A055 modificado	186	-	+
240B10D	Cry3A055 modificado	104	-	+
240B10E	Cry3A055 modificado	70	-	+
240A15E	Cry3A055 modificado	122	-	+
240B4D	Cry3A055 modificado	97	-	+
243B5A	Cry3A055 modificado	41	-	+

244A7A	Cry3A055 modificado	191	-	+
710-2-51	Maíz optimizado	39	+	-
710-2-54	Maíz optimizado	857	+	-
710-2-61	Maíz optimizado	241	+	-
710-2-67	Maíz optimizado	1169	+	-
710-2-68	Maíz optimizado	531	+	-
710-2-79	Maíz optimizado	497	+	-
710-2-79	Maíz optimizado	268	+	-
WT1 Control	-	0	+	-
WT2 Control	-	0	+	-

REIVINDICACIONES

1. Una molécula de ácido nucleico aislada que comprende una secuencia nucleotídica que codifica una toxina Cry3A modificada que comprende un sitio de reconocimiento de proteasa no natural, donde dicho sitio de reconocimiento de proteasa modifica una toxina Cry3A y está situada en una posición seleccionada del grupo que consiste en:

a) una posición comprendida entre los aminoácidos que corresponden a los aminoácidos números 107 y 115 de la SEC ID NO: 4;

b) una posición comprendida entre los aminoácidos que corresponden a los aminoácidos números 536 y 542 de la SEC ID NO: 4; y

c) una posición comprendida entre los aminoácidos que corresponden a los aminoácidos números 107 y 115 de la SEC ID NO: 4; y entre los aminoácidos que corresponden a los aminoácidos números 536 y 542 de la SEC ID NO: 4,

donde dicho sitio de reconocimiento de proteasa reconocible por una proteasa intestinal del gusano de la raíz del maíz del oeste, y donde dicha toxina Cry3A modificada causa una mortalidad más elevada en el gusano de la raíz del maíz del oeste que la mortalidad causada por dichas toxinas Cry3A al gusano de la raíz del maíz del oeste en un bioensayo de dieta artificial.

2. La molécula de ácido nucleico aislada de acuerdo con la reivindicación 1, donde dicha protéasa intestinal es catepsina G.

3. La molécula de ácido nucleico aislada de acuerdo con la reivindicación 1, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos número: 107 y 115 de la SEC ID NO: 4.

4. La molécula de ácido nucleico aislada de acuerdo con la reivindicación 1, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos que corresponden a los aminoácidos número: 107 y 113 de la SEC ID NO: 4.

5. La molécula de ácido nucleico aislada de acuerdo con la reivindicación 4, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos número: 107 y 113 de la SEC ID NO: 4.

6. La molécula de ácido nucleico aislada de acuerdo con la reivindicación 1, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos que corresponden a los aminoácidos número: 107 y 111 de la SEC ID NO: 4.

7. La molécula de ácido nucleico aislada de acuerdo con la reivindicación 6, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos número: 107 y 111 de la SEC ID NO: 4.

8. La molécula de ácido nucleico aislada de acuerdo con la reivindicación 1, donde dicho sitio de proteasa está situado entre los aminoácidos número: 536 y 542 de la SEC ID NO: 4.

9. La molécula de ácido nucleico aislada de acuerdo con la reivindicación 1, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos que corresponden a los aminoácidos número: 536 y 541 de la SEC ID NO: 4.

10. La molécula de ácido nucleico aislada de acuerdo con la reivindicación 9, donde dicho sitio de proteasa adicional está situado entre los aminoácidos número: 536 y 541 de la SEC ID NO: 4.

11. La molécula de ácido nucleico aislada de acuerdo con la reivindicación 1, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos que corresponden a los aminoácidos números: 540 y 541 de la SEC ID NO: 4.

12. La molécula de ácido nucleico aislada de acuerdo con la reivindicación 11, donde dicho sitio de proteasa está situado entre los aminoácidos números: 540 y 541 de la SEC ID NO: 4.

13. La molécula de ácido nucleico aislada de acuerdo con la reivindicación 1, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos números: 107 y 115 de la SEC ID NO: 4, y entre los aminoácidos números: 536 y 542 de la SEC ID NO. 4.

14. La molécula de ácido nucleico aislada de acuerdo con la reivindicación 1, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos que corresponden a los aminoácidos números: 107 y 113 de la SEC ID NO: 4 y entre los aminoácidos que corresponden a los aminoácidos números: 540 y 541 de la SEC ID NO: 4.

15. La molécula de ácido nucleico aislada de acuerdo con la reivindicación 14, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos números 107 y 113 de la SEC ID NO: 4 y entre los aminoácidos números 540 y 541 de la SEC ID NO: 4.

16. La molécula de ácido nucleico aislada de acuerdo con la reivindicación 1, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos que corresponden a los aminoácidos números: 107 y 111 de la SEC ID NO: 4 y entre los aminoácidos que corresponden a los aminoácidos números: 536 y 541 de la SEC ID NO: 4.

17. La molécula de ácido nucleico aislada de acuerdo con la reivindicación 16, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos números 107 y 111 de la SEC ID NO: 4 y entre los aminoácidos números 536 y 541 de la SEC ID NO: 4.

18. La molécula de ácido nucleico aislada de acuerdo con la reivindicación 1, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos que corresponden a los aminoácidos números: 107 y 111 de la SEC ID NO: 4 y entre los aminoácidos que corresponden a los aminoácidos números: 540 y 541 de la SEC ID NO: 4.

19. La molécula de ácido nucleico aislada de acuerdo con la reivindicación 18, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos números 107 y 111 de la SEC ID NO: 4 y entre los aminoácidos números 540 y 541 de la SEC ID NO: 4.

20. La molécula de ácido nucleico aislada de acuerdo con la reivindicación 1, donde dicha toxina Cry3A modificada causa una mortalidad de por lo menos el 50% en el gusano de la raíz del maíz del oeste en el cual dicha toxina Cry3A causa hasta una 30% de mortalidad.

21. La molécula de ácido nucleico aislada de acuerdo con la reivindicación 1, donde dicha secuencia nucleotídica comprende los nucleótidos 1-1791 de la SEC ID NO: 6, los nucleótidos 1-1806 de la SEC ID NO: 8, los nucleótidos 1-1812 de la SEC ID NO: 10, los nucleótidos 1-1794 de la SEC ID NO: 12, los nucleótidos 1-1818 de la SEC ID NO: 14, los nucleótidos 1-1812 de la SEC ID NO: 16, los nucleótidos 1-1791 de la SEC ID NO: 18 o los nucleótidos 1-1818 de la SEC ID NO: 20.

22. La molécula de ácido nucleico aislada de acuerdo con la reivindicación 1, donde dicha toxina Cry3A modificada comprende la secuencia de aminoácido establecida en las SEC ID NO: 7, SEC ID NO: 9, SEC ID NO: 11, SEC ID NO: 13, SEC ID NO: 15, SEC ID NO: 17, SEC ID NO: 19, SEC ID NO: 21.

23. La molécula de ácido nucleico aislada de acuerdo con la reivindicación 1, donde dicha toxina Cry3A modificada es activa contra el gusano de la raíz del maíz del norte.

24. Una construcción quimérica que comprende una secuencia promotora heteróloga ligada operativamente a la molécula de ácido nucleico de la reivindicación 1.

25. Un vector recombinante que comprende la construcción quimérica de la reivindicación 24.

26. Una célula huésped no humana transgénica que comprende la construcción quimérica de la reivindicación 24.

27. La célula huésped transgénica de acuerdo con la reivindicación 26 que es una célula bacteriana.

28. La célula huésped transgénica de acuerdo con la reivindicación 26 que es una célula de planta.

29. Una planta transgénica que comprende la célula de planta transgénica de la reivindicación 28.

30. La planta transgénica de acuerdo con la reivindicación 29, donde dicha planta es una planta de maíz.

31. La semilla transgénica de la planta transgénica de la reivindicación 29.

32. La semilla transgénica de la planta de maíz de la reivindicación 30.

33. Una toxina aislada producida por la expresión de la molécula de ácido nucleico de acuerdo con la reivindicación 1.

34. La toxina aislada de acuerdo con la reivindicación 33, donde dicha toxina se produce por la expresión de una molécula de ácido nucleico que comprende los nucleótidos 1-1791 de la SEC ID NO: 6, los nucleótidos 1-1806 de la SEC ID NO: 8, los nucleótidos 1-1812 de la SEC ID NO: 10, los nucleótidos 1-1794 de la SEC ID NO: 12, los nucleótidos 1-1818 de la SEC ID NO: 14, los nucleótidos 1-1812 de la SEC ID NO: 16, los nucleótidos 1-1791 de la SEC ID NO: 18 o los nucleótidos 1-1818 de la SEC ID NO: 20.

35. La toxina aislada de acuerdo con la reivindicación 33, donde dicha toxina tiene actividad contra el gusano de la raíz del maíz del norte.

36. La toxina aislada de acuerdo con la reivindicación 33, donde dicha toxina comprende la secuencia de aminoácido establecida en las SEC ID NO: 7, SEC ID NO: 9, SEC ID NO: 11, SEC ID NO: 13, SEC ID NO: 15, SEC ID NO: 17, SEC ID NO: 19, SEC ID NO: 21.

37. Una composición que comprende una cantidad eficaz de la toxina de la reivindicación 33, que causa mortalidad al gusano de la raíz del maíz del oeste.

38. Una planta transgénica de maíz que comprende una secuencia nucleotídica que codifica una toxina Cry3A modificada que comprende un sitio de reconocimiento de proteasa no natural, donde dicho sitio de reconocimiento de proteasa modifica una toxina Cry3A y está situada en una posición seleccionada del grupo que consiste en:

a) una posición comprendida entre los aminoácidos que corresponden a los aminoácidos números 107 y 115 de la SEC ID NO: 4;

b) una posición comprendida entre los aminoácidos que corresponden a los aminoácidos números 536 y 542 de la SEC ID NO: 4; y

c) una posición comprendida entre los aminoácidos que corresponden a los aminoácidos números 107 y 115 de la SEC ID NO: 4; y entre los aminoácidos que corresponden a los aminoácidos números 536 y 542 de la SEC ID NO: 4,

donde dicho sitio de reconocimiento de proteasa reconocible por una proteasa intestinal del gusano de la raíz del maíz del oeste, y donde dicha planta transgénica expresa dicha toxina Cry3A modificada en un tejido de raíz a un nivel que causa mortalidad en el gusano de la raíz del maíz del oeste.

39. La planta transgénica del maíz de acuerdo con la reivindicación 38, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos números 107 y 111 de la SEC ID NO: 4.

40. La planta transgénica del maíz de acuerdo con la reivindicación 38, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos números 540 y 541 de la SEC ID NO: 4.

41. La planta transgénica del maíz de acuerdo con la reivindicación 38, donde dicho sitio de reconocimiento de proteasa está situado entre los

aminoácidos números 107 y 111 de la SEC ID NO: 4 y entre los aminoácidos números 540 y 541 de la SEC ID NO: 4.

42. La planta transgénica del maíz de acuerdo con la reivindicación 38, donde dicha secuencia nucleotídica comprende los nucleótidos 1-1806 de la SEC ID NO: 8, los nucleótidos 1-1818 de la SEC ID NO: 14, o los nucleótidos 1-1791 de la SEC ID NO: 18.

43. La planta transgénica de maíz de acuerdo con la reivindicación 38 donde dicha toxina Cry3A modificada comprende la secuencia de aminoácido establecida en la SEC ID NO: 9, SEC ID NO: 15 O SEC ID NO: 19.

44. La planta transgénica de maíz de acuerdo con la reivindicación 38 donde dicho tejido de raíz causa 100 % de mortalidad en el gusano de la raíz del maíz del oeste.

45. La planta transgénica de maíz de acuerdo con la reivindicación 38 donde dicho tejido de raíz causa 90 % de mortalidad en el gusano de la raíz del maíz del oeste.

46. La planta transgénica de maíz de acuerdo con la reivindicación 38 donde dicho tejido de raíz causa 80 % de mortalidad en el gusano de la raíz del maíz del oeste.

47. La planta transgénica de maíz de acuerdo con la reivindicación 38 donde dicho tejido de raíz causa 70 % de mortalidad en el gusano de la raíz del maíz del oeste.

48. La planta transgénica de maíz de acuerdo con la reivindicación 38 donde dicho tejido de raíz causa 60 % de mortalidad en el gusano de la raíz del maíz del oeste.

49. La planta transgénica de maíz de acuerdo con la reivindicación 38 donde dicho tejido de raíz causa 50 % de mortalidad en el gusano de la raíz del maíz del oeste.

50. La planta transgénica de maíz de acuerdo con la reivindicación 38 donde dicho tejido de raíz causa 40 % de mortalidad en el gusano de la raíz del maíz del oeste.

51. La planta transgénica de maíz de acuerdo con la reivindicación 38 donde dicha planta transgénica expresa dicha toxina Cry3A modificada a un nivel suficiente para evitar que el gusano de la raíz del maíz del oeste cercene gravemente las raíces de la planta transgénica.

52. La planta transgénica de maíz de acuerdo con la reivindicación 38 donde dicha planta transgénica expresa dicha toxina Cry3A modificada a un nivel suficiente para evitar el perjuicio por alimentación del gusano de la raíz del maíz del oeste que causa el debilitamiento de la planta.

53. La planta transgénica de maíz de acuerdo con cualquiera de las reivindicaciones 38-52, que es una planta endogámica.

54. La planta transgénica de acuerdo con cualquiera de las reivindicación 38-52, que es una planta híbrida.

55. Semilla transgénica de la planta de la reivindicación 53.

56. Semilla transgénica de la planta de la reivindicación 54.

57. Una toxina Cry3A modificada que comprende un sitio de reconocimiento de proteasa no natural, donde dicho sitio de reconocimiento de proteasa modifica una toxina Cry3A y está situado en una posición seleccionada del grupo que consiste en:

a) una posición comprendida entre los aminoácidos que corresponden a los aminoácidos números 107 y 115 de la SEC ID NO: 4;

b) una posición comprendida entre los aminoácidos que corresponden a los aminoácidos números 536 y 542 de la SEC ID NO: 4; y

c) una posición comprendida entre los aminoácidos que corresponden a los aminoácidos números 107 y 115 de la SEC ID NO: 4; y entre los aminoácidos que corresponden a los aminoácidos números 536 y 542 de la SEC ID NO: 4,
donde dicho sitio de reconocimiento de proteasa es reconocible por una proteasa intestinal del gusano de la raíz del maíz del oeste, y donde dicha toxina Cry3A modificada causa una mortalidad más elevada en el gusano de la raíz del maíz del oeste que la mortalidad causada por dichas toxinas Cry3A al gusano de la raíz del maíz del oeste en un bioensayo de dieta artificial.

58. La toxina Cry3A modificada de acuerdo con la reivindicación 57, donde dicha proteasa intestinal es catepsina G.

59. La toxina Cry3A modificada de acuerdo con la reivindicación 57, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos 107 y 115 de la SEC ID NO: 4.

60. La toxina Cry3A modificada de acuerdo con la reivindicación 57, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos que corresponden a los aminoácidos números 107 y 113 de la SEC ID NO: 4.

61. La toxina Cry3A modificada de acuerdo con la reivindicación 60, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos números 107 y 113 de la SEC ID NO: 4.

62. La toxina Cry3A modificada de acuerdo con la reivindicación 57, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos que corresponden a los aminoácidos números 107 y 111 de la SEC ID NO: 4.

63. La toxina Cry3A modificada de acuerdo con la reivindicación 62, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos números 107 y 111 de la SEC ID NO: 4.

64. La toxina Cry3A modificada de acuerdo con la reivindicación 57, donde dicho sitio de proteasa está situado entre los aminoácidos números 536 y 542 de la SEC ID NO: 4.

65. La toxina Cry3A modificada de acuerdo con la reivindicación 57, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos que corresponden a los aminoácidos números 536 y 541 de la SEC ID NO: 4.

66. La toxina Cry3A modificada de acuerdo con la reivindicación 65, donde dicho sitio de reconocimiento de proteasa adicional está situado entre los aminoácidos números 536 y 541 de la SEC ID NO: 4.

67. La toxina Cry3A modificada de acuerdo con la reivindicación 57, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos que corresponden a los aminoácidos números 540 y 541 de la SEC ID NO: 4.

68. La toxina Cry3A modificada de acuerdo con la reivindicación 57, donde dicho sitio de proteasa está situado entre los aminoácidos números 540 y 541 de la SEC ID NO: 4.

69. La toxina Cry3A modificada de acuerdo con la reivindicación 57, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos números 107 y 115 y entre los aminoácidos números 536 y 542 de la SEC ID NO: 4.

70. La toxina Cry3A modificada de acuerdo con la reivindicación 57, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos que corresponden a los aminoácidos números 107 y 113 de la SEC ID NO: 4 y entre los aminoácidos que corresponden a los aminoácidos números 540 y 541 de la SEC ID NO: 4.

71. La toxina Cry3A modificada de acuerdo con la reivindicación 70, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos 107

y 113 de la SEC ID NO: 4 y entre los aminoácidos números 540 y 541 de la SEC ID NO: 4.

72. La toxina Cry3A modificada de acuerdo con la reivindicación 57, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos que corresponden a los aminoácidos números 107 y 111 de la SEC ID NO: 4 y entre los aminoácidos que corresponden a los aminoácidos números 536 y 541 de la SEC ID NO: 4.

73. La toxina Cry3A modificada de acuerdo con la reivindicación 72, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos números 107 y 111 de la SEC ID NO: 4 y entre los aminoácidos números 536 y 541 de la SEC ID NO: 4.

74. La toxina Cry3A modificada de acuerdo con la reivindicación 57, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos que corresponden a los aminoácidos números 107 y 111 de la SEC ID NO: 4 y entre los aminoácidos que corresponden a los aminoácidos números 540 y 541 de la SEC ID NO: 4.

75. La toxina Cry3A modificada de acuerdo con la reivindicación 74, donde dicho sitio de reconocimiento de proteasa está situado entre los aminoácidos números 107 y 111 de la SEC ID NO: 4 y entre los aminoácidos que corresponden a los aminoácidos números 540 y 541 de la SEC ID NO: 4.

76. La toxina Cry3A modificada de acuerdo con la reivindicación 57, donde dicha toxina cry3A modificada causa por lo menos 50% de mortalidad en el gusano de la raíz del maíz del oeste en el cual dicha toxina Cry3A causa una mortalidad de hasta 30%.

77. La toxina Cry3A modificada de acuerdo con la reivindicación 57 que es codificada por los nucleótidos 1-1791 de la SEC ID NO: 6, los nucleótidos 1-1806 de la SEC ID NO: 8, los nucleótidos 1-1812 de la SEC ID NO: 10, los nucleótidos

1-1794 de la SEC ID NO: 12, los nucleótidos 1-1818 de la SEC ID NO: 14, los nucleótidos 1-1812 de la SEC ID NO: 16, los nucleótidos 1-1791 de la SEC ID NO: 18 o los nucleótidos 1-1818 de la SEC ID NO: 20.

78. La toxina Cry3a modificada de acuerdo con la reivindicación 57 que comprende la secuencia de aminoácido establecida en las SEC ID NO: 7, SEC ID NO: 9, SEC ID NO: 11, SEC ID NO: 13, SEC ID NO: 15, SEC ID NO: 17, SEC ID NO: 19, SEC ID NO: 21.

79. La toxina Cry3a modificada de acuerdo con la reivindicación 57 que es activa contra el gusano de la raíz del maíz del norte.

80. Un método para controlar la infestación de plantas de maíz del gusano de la raíz del maíz del oeste, donde dicho método comprende:

a) proveer la planta transgénica de maíz de acuerdo con la reivindicación 38; y

b) poner en contacto dicho gusano de la raíz del maíz del oeste con la planta.

81. Un método para producir una toxina Cry3A modificada que comprende:

a) obtener la célula huésped transgénica de acuerdo con la reivindicación 27;

b) cultivar la célula huésped transgénica bajo condiciones que permiten la expresión de la toxinas Cry3A modificadas; y

c) recuperar la toxina Cry3A modificada expresada.

82. Un método para producir plantas resistentes a los insectos, que comprende:

(a) integrar establemente la molécula de ácido nucleico de acuerdo con la reivindicación 1, en el genoma de células de plantas, y

(b) regenerar las plantas establemente transformadas de dichas células de plantas transformadas, donde dichas plantas establemente transformadas expresan una cantidad eficaz de una toxina Cry3A modificada que convierte dicha planta transformada en resistente a por lo menos el gusano de la raíz del maíz del oeste.

83. Un método para controlar por lo menos al gusano de la raíz del maíz del oeste, que comprende liberar oralmente al gusano de la raíz del maíz del oeste una cantidad eficaz de una toxina de acuerdo con la reivindicación 33.

84. Un método para preparar un toxina Cry3A modificada, que comprende:

(a) obtener un gen *cry*3A que codifica una toxina Cry3A;

(b) identificar una proteasa intestinal del gusano de la raíz del maíz del oeste;

(c) obtener una secuencia nucleotídica que codifica un sitio de reconocimiento para dicha proteasa intestinal;

(d) insertar dicha secuencia nucleotídica en dicho gen *cry*3A, de modo que dicho sitio de reconocimiento esté situado en dicha toxina Cry3A en una posición comprendida entre los aminoácidos que corresponden a los aminoácidos números 107 y 115 de la SEC ID NO: 4, o en una posición comprendida entre los aminoácidos que corresponden a los aminoácidos números 536 y 542 de la SEC ID NO: 4, o en una posición comprendida entre los aminoácidos que corresponden a los aminoácidos números 107 y 115 de la SEC ID NO: 4, y entre los aminoácidos que corresponden a los aminoácidos números 536 y 542 de la SEC ID NO: 4, creando de este modo el gen *cry*3A modificado;

(e) insertar dicho gen *cry*3A modificado en un cassette de expresión; y

(f) transformar dicho cassette de expresión en una célula huésped no humana, donde dicha célula huésped produce una toxina Cry3A modificada.

85. Un gen *cry*3A modificado que comprende una secuencia nucleotídica que codifica una toxina Cry3A modificada que comprende un sitio de reconocimiento de proteasa que no es natural, donde dicho gen *cry*3A modificado comprende una secuencia codificadora que codifica dicho sitio de reconocimiento de proteasa donde dicha secuencia codificadora modifica un gen *cry*3A y está insertado en una posición seleccionada del grupo que consiste en:

a) una posición comprendida entre los codones que codifican para los aminoácidos que corresponden a los aminoácidos números 107 y 115 de la SEC ID NO: 4;

b) una posición entre los codones que codifican para aminoácidos que corresponden a los aminoácidos números 536 y 542 de la SEC ID NO: 4; y

c) una posición comprendida entre los codones que codifican para aminoácidos que corresponden a los aminoácidos números 107 y 115 de la SEC ID NO: 4; y entre codones que codifican para aminoácidos que corresponden a los aminoácidos números 536 y 542 de la SEC ID NO: 4,

donde dicho sitio de reconocimiento de proteasa es reconocible por una proteasa intestinal del gusano de la raíz del maíz del oeste, y donde dicha toxina Cry3A modificada causa una mortalidad más elevada en el gusano de la raíz del maíz del oeste que la mortalidad causada por dichas toxinas Cry3A al gusano de la raíz del maíz del oeste en un bioensayo de dieta artificial.

86. El gen *cry*3A modificado de acuerdo con la reivindicación 85, donde dicha proteasa intestinal es catepsina G.

87. El gen *cry*3A modificado de acuerdo con la reivindicación 85 donde dicha secuencia nucleotídica comprende los nucleótidos 1-1791 de la SEC ID NO: 6, los nucleótidos 1-1806 de la SEC ID NO: 8, los nucleótidos 1-1812 de la SEC ID NO: 10, los nucleótidos 1-1794 de la SEC ID NO: 12, los nucleótidos 1-1818 de la SEC ID NO: 14, los nucleótidos 1-1812 de la SEC ID NO: 16, los nucleótidos 1-1791 de la SEC ID NO: 18 o los nucleótidos 1-1818 de la SEC ID NO: 20.

88. El gen *cry*3a modificado de acuerdo con la reivindicación 85 donde dicha toxina Cry3A modificada comprende la secuencia de aminoácido establecida en las SEC ID NO: 7, SEC ID NO: 9, SEC ID NO: 11, SEC ID NO: 13, SEC ID NO: 15, SEC ID NO: 17, SEC ID NO: 19, SEC ID NO: 21.

89. El gen *cry*3A modificado de acuerdo con la reivindicación 85 donde dicha toxina Cry3A modificada es activa contra el gusano de la raíz del maíz del norte.

90. Una construcción quimérica que comprende una secuencia promotora heteróloga ligada operativamente al gen *cry*3A modificado de la reivindicación 85.

91. Un vector recombinante que comprende la construcción quimérica de la reivindicación 90.

92. Una célula huésped no humana transgénica que comprende la construcción quimérica de la reivindicación 90.

93. Una célula huésped transgénica de acuerdo con la reivindicación 92 que es una célula bacteriana.

94. La célula huésped transgénica de acuerdo con la reivindicación 92 que es una célula de planta.

95. Una planta transgénica que comprende la célula de planta transgénica de la reivindicación 94.

96. La planta transgénica de acuerdo con la reivindicación 95, donde dicha planta es una planta de maíz.

97. La semilla transgénica de la planta transgénica de la reivindicación 95.

98. La semilla transgénica de la planta de maíz de la reivindicación 96.

99. Una planta de maíz transgénica que comprende un gen *cry*3A modificado que comprende una secuencia nucleotídica que codifica la toxina Cry3A modificada que comprende un sitio de reconocimiento de proteasa no natural, donde dicho gen *cry*3A modificado comprende una secuencia codificadora que codifica dicho sitio de reconocimiento de proteasa donde dicha secuencia codificadora modifica un gen *cry*3A y está insertado en una posición seleccionada del grupo que consiste en:

a) una posición comprendida entre los codones que codifican para los aminoácidos que corresponden a los aminoácidos números 107 y 115 de la SEC ID NO: 4;

b) una posición entre los codones que codifican para aminoácidos que corresponden a los aminoácidos números 536 y 542 de la SEC ID NO: 4; y

c) una posición comprendida entre los codones que codifican para aminoácidos que corresponden a los aminoácidos números 107 y 115 de la SEC ID NO: 4; y entre codones que codifican para aminoácidos que corresponden a los aminoácidos números 536 y 542 de la SEC ID NO: 4,

donde dicho sitio de reconocimiento de proteasa es reconocible por una proteasa intestinal del gusano de la raíz del maíz del oeste, y donde dicha planta transgénica expresa dicha toxina Cry3A modificada en el tejido de la raíz a un nivel que causa mortalidad al gusano de la raíz del maíz del oeste.

100. La planta de maíz transgénica de la reivindicación 99, donde dicha secuencia nucleotídica comprende los nucleótidos 1-1806 de la SEC ID NO: 8, nucléotidos 1-1818 de la SEC ID NO: 14, o nucleóetidos 1-1791 de la SEC ID NO: 18.

101. La planta de maíz transgénica de la reivindicación 99, donde dicha toxina Cry3A modificada comprende la secuencia de aminoácido establecida en la SEC ID NO: 9, SEC ID NO: 15, O SEC ID NO: 19.

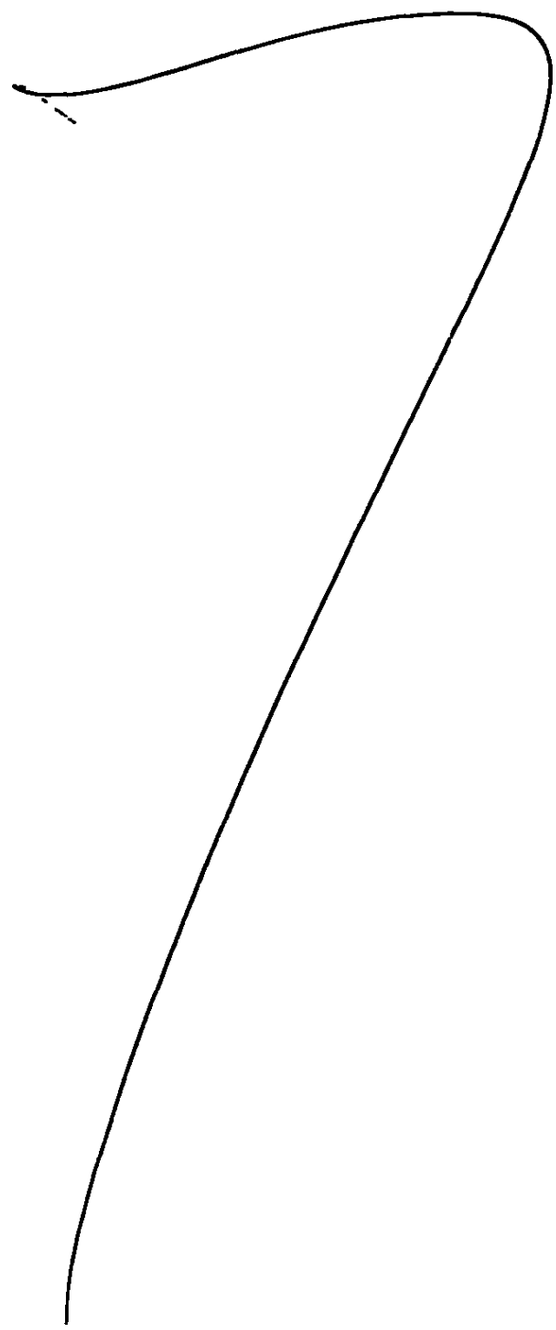
102. La planta de maíz transgénica de acuerdo con cualquiera de las reivindicaciones 99-101, que es una planta endogámica.

103. La planta de maíz transgénica de acuerdo con cualquiera de las reivindicaciones 99-101, que es una planta híbrida.

190
192

104. Semilla transgénica de la planta de la reivindicación 102.

105. Semilla transgénica de la planta de la reivindicación 103.



RESUMEN

Se describen composiciones y métodos para controlar plagas de plantas. En particular, se proveen nuevas secuencias de ácido nucleico que codifican toxina Cry3A modificadas que tienen toxicidad incrementada para el gusano de la raíz del maíz. Mediante la inserción de un sitio de reconocimiento de proteasa, tal como Catepsina G, que es reconocida por una proteasa intestinal de un insecto objetivo, en por lo menos una posición de una toxina Cry3A, se diseña una toxina Cry3A modificada que tiene una toxicidad significativamente mayor, particularmente para el gusano de la raíz del maíz del oeste y del norte. Además, se describe un método para preparar las toxinas Cry3A modificadas y métodos de uso de las secuencias de ácido nucleico Cry3A, por ejemplo microorganismos para el control de insectos o en plantas transgénicas para conferir protección contra el daño causado por insectos, y un método de uso de las toxinas Cry3A modificadas y composiciones y formulaciones que comprenden las toxinas Cry3A, por ejemplo, aplicar las toxinas Cry3A o las composiciones o formulaciones a áreas infestadas de insectos o para tratar profilácticamente áreas o plantas susceptibles a los insectos con el propósito de conferir protección contra las plagas de insectos.

LISTADO DE SECUENCIAS

<110> Participaciones Syngenta AG

<120> Toxinas Cry3A Modificadas y Secuencias de Acido Nucleico que las Codifican
<130> 60065/PCT

<140>
<141>

<150> US 60/316421
<151> 2001-08-31

<160> 34

<170> PatentIn Ver. 3.0

<210> 1
<211> 1932
<212> ADN
<213> Bacillus thuringiensis

<220>
<221> CDS
<222> (1)..(1932)
<223> Secuencia codificadora de cry3A natural de acuerdo con Sekar et al. 1987, Proc. Natl. Acad. Sci. 84:7036-7040

<400> 1
atg aat ccg aac aat cga agt gaa cat gat aca ata aaa act act gaa 48
Met Asn Pro Asn Asn Arg Ser Glu His Asp Thr Ile Lys Thr Thr Glu
1 5 10 15

aat aat gag gtg cca act aac cat gtt caa tat cct tta gcg gaa act 96
Asn Asn Glu Val Pro Thr Asn His Val Gln Tyr Pro Leu Ala Glu Thr
20 25 30

cca aat cca aca cta gaa gat tta aat tat aaa gag ttt tta aga atg 144
Pro Asn Pro Thr Leu Glu Asp Leu Asn Tyr Lys Glu Phe Leu Arg Met
35 40 45

act gca gat aat aat acg gaa gca cta gat agc tct aca aca aaa gat 192
Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys Asp
50 55 60

gtc att caa aaa ggc att tcc gta gta ggt gat ctc cta ggc gta gta 240
Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val Val
65 70 75 80

ggt ttc ccg ttt ggt gga gcg ctt gtt tcg ttt tat aca aac ttt tta 288
Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe Leu
85 90 95

aat act att tgg cca agt gaa gac ccg tgg aag gct ttt atg gaa caa 336
Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu Gln
100 105 110

gta gaa gca ttg atg gat cag aaa ata gct gat tat gca aaa aat aaa 384
Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn Lys
115 120 125

gct ctt gca gag tta cag ggc ctt caa aat aat gtc gaa gat tat gtg 432
Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr Val
130 135 140

agt gca ttg agt tca tgg caa aaa aat cct gtg agt tca cga aat cca 480
Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn Pro
145 150 155 160

cat agc cag ggg cgg ata aga gag ctg ttt tct caa gca gaa agt cat 528
His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser His
165 170 175

ttt cgt aat tca atg cct tcg ttt gca att tct gga tac gag gtt cta 576
Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val Leu
180 185 190

ttt cta aca aca tat gca caa gct gcc aac aca cat tta ttt tta cta 624
Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu Leu
195 200 205

aaa gac gct caa att tat gga gaa gaa tgg gga tac gaa aaa gaa gat 672
Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu Asp
210 215 220

att gct gaa ttt tat aaa aga caa cta aaa ctt acg caa gaa tat act 720
Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr Thr
225 230 235 240

gac cat tgt gtc aaa tgg tat aat gtt gga tta gat aaa tta aga ggt 768
Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg Gly
245 250 255

tca tct tat gaa tct tgg gta aac ttt aac cgt tat cgc aga gag atg 816
Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu Met
260 265 270

aca tta aca gta tta gat tta att gca cta ttt cca ttg tat gat gtt 864
Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp Val
275 280 285

cgg cta tac cca aaa gaa gtt aaa acc gaa tta aca aga gac gtt tta 912
Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val Leu
290 295 300

aca gat cca att gtc gga gtc aac aac ctt agg ggc tat gga aca acc 960
Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr Thr
305 310 315 320

ttc tct aat ata gaa aat tat att cga aaa cca cat cta ttt gac tat 1008
Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp Tyr
325 330 335

ctg cat aga att caa ttt cac acg cgg ttc caa cca gga tat tat gga 1056
Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr Gly
340 345 350

aat gac tct ttc aat tat tgg tcc ggt aat tat gtt tca act aga cca 1104
Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg Pro
355 360 365

agc ata gga tca aat gat ata atc aca tct cca ttc tat gga aat aaa 1152
Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn Lys
370 375 380

tcc agt gaa cct gta caa aat tta gaa ttt aat gga gaa aaa gtc tat 1200
Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val Tyr
385 390 395 400

aga gcc gta gca aat aca aat ctt gcg gtc tgg ccg tcc gct gta tat 1248
Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val Tyr
405 410 415

tca ggt gtt aca aaa gtg gaa ttt agc caa tat aat gat caa aca gat 1296
Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr Asp
420 425 430

gaa gca agt aca caa acg tac gac tca aaa aga aat gtt ggc gcg gtc 1344
Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala Val
435 440 445

agc tgg gat tct atc gat caa ttg cct cca gaa aca aca gat gaa cct 1392
Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu Pro
450 455 460

cta gaa aag gga tat agc cat caa ctc aat tat gta atg tgc ttt tta 1440
Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe Leu
465 470 475 480

atg cag ggt agt aga gga aca atc cca gtg tta act tgg aca cat aaa 1488
Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His Lys
485 490 495

agt gta gac ttt ttt aac atg att gat tcg aaa aaa att aca caa ctt 1536
Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln Leu
500 505 510

ccg tta gta aag gca tat aag tta caa tct ggt gct tcc gtt gtc gca 1584
Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val Ala
515 520 525

ggt cct agg ttt aca gga gga gat atc att caa tgc aca gaa aat gga 1632
Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn Gly
530 535 540

agt gcg gca act att tac gtt aca ccg gat gtg tcg tac tct caa aaa 1680
Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln Lys
545 550 555 560

tat cga gct aga att cat tat gct tct aca tct cag ata aca ttt aca 1728
Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe Thr
565 570 575

ctc agt tta gac ggg gca cca ttt aat caa tac tat ttc gat aaa acg 1776
Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp Lys Thr
580 585 590

ata aat aaa gga gac aca tta acg tat aat tca ttt aat tta gca agt 1824
Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala Ser
595 600 605

ttc agc aca cca ttc gaa tta tca ggg aat aac tta caa ata ggc gtc 1872
Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly Val
610 615 620

aca gga tta agt gct gga gat aaa gtt tat ata gac aaa att gaa ttt 1920
Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu Phe
625 630 635 640

att cca gtg aat 1932
Ile Pro Val Asn

<210> 2
<211> 644
<212> PRT
<213> Bacillus thuringiensis

<400> 2
Met Asn Pro Asn Asn Arg Ser Glu His Asp Thr Ile Lys Thr Thr Glu

~~197~~
199

1 5 10 15

Asn Asn Glu Val Pro Thr Asn His Val Gln Tyr Pro Leu Ala Glu Thr
20 25 30

Pro Asn Pro Thr Leu Glu Asp Leu Asn Tyr Lys Glu Phe Leu Arg Met
35 40 45

Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys Asp
50 55 60

Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val Val
65 70 75 80

Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe Leu
85 90 95

Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu Gln
100 105 110

Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn Lys
115 120 125

Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr Val
130 135 140

Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn Pro
145 150 155 160

His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser His
165 170 175

Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val Leu
180 185 190

Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu Leu
195 200 205

Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu Asp
210 215 220

~~198~~
100

Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr Thr
225 230 235 240

Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg Gly
245 250 255

Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu Met
260 265 270

Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp Val
275 280 285

Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val Leu
290 295 300

Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr Thr
305 310 315 320

Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp Tyr
325 330 335

Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr Gly
340 345 350

Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg Pro
355 360 365

Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn Lys
370 375 380

Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val Tyr
385 390 395 400

Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val Tyr
405 410 415

Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr Asp
420 425 430

Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala Val
435 440 445

Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu Pro
450 455 460

Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe Leu
465 470 475 480

Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His Lys
485 490 495

Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln Leu
500 505 510

Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val Ala
515 520 525

Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn Gly
530 535 540

Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln Lys
545 550 555 560

Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe Thr
565 570 575

Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp Lys Thr
580 585 590

Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala Ser
595 600 605

Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly Val
610 615 620

Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu Phe
625 630 635 640

Ile Pro Val Asn

<210> 3
<211> 1803
<212> ADN
<213> Secuencia Artificial

<220>
<221> CDS
<222> (1)..(1794)
<223> Secuencia codificadora de cry3A de maíz optimizado

<400> 3
atg acg gcc gac aac aac acc gag gcc ctg gac agc agc acc acc aag 48
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

gac gtg atc cag aag ggc atc agc gtg gtg ggc gac ctg ctg ggc gtg 96
Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

gtg ggc ttc ccc ttc ggc ggc gcc ctg gtg agc ttc tac acc aac ttc 144
Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

ctg aac acc atc tgg ccc agc gag gac ccc tgg aag gcc ttc atg gag 192
Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

cag gtg gag gcc ctg atg gac cag aag atc gcc gac tac gcc aag aac 240
Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65 70 75 80

aag gca ctg gcc gag cta cag ggc ctc cag aac aac gtg gag gac tat 288
Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

gtg agc gcc ctg agc agc tgg cag aag aac ccc gtc tcg agc cgc aac 336
Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn
100 105 110

ccc cac agc cag ggc cgc atc cgc gag ctg ttc agc cag gcc gag agc 384
Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser
115 120 125

cac ttc cgc aac agc atg ccc agc ttc gcc atc agc ggc tac gag gtg 432
His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val
130 135 140

ctg ttc ctg acc acc tac gcc cag gcc gcc aac acc cac ctg ttc ctg 480
Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu
145 150 155 160

ctg aag gac gcc caa atc tac gga gag gag tgg ggc tac gag aag gag 528
Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu
165 170 175

gac atc gcc gag ttc tac aag cgc cag ctg aag ctg acc cag gag tac 576
Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr
180 185 190

acc gac cac tgc gtg aag tgg tac aac gtg ggt cta gac aag ctc cgc 624
Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg
195 200 205

ggc agc agc tac gag agc tgg gtg aac ttc aac cgc tac cgc cgc gag 672
Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu
210 215 220

atg acc ctg acc gtg ctg gac ctg atc gcc ctg ttc ccc ctg tac gac 720
Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp
225 230 235 240

gtg cgc ctg tac ccc aag gag gtg aag acc gag ctg acc cgc gac gtg 768
Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val
245 250 255

ctg acc gac ccc atc gtg ggc gtg aac aac ctg cgc ggc tac ggc acc 816
Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr
260 265 270

~~202~~
204

acc ttc agc aac atc gag aac tac atc cgc aag ccc cac ctg ttc gac 864
Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp
275 280 285

tac ctg cac cgc atc cag ttc cac acg cgt ttc cag ccc ggc tac tac 912
Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr
290 295 300

ggc aac gac agc ttc aac tac tgg agc ggc aac tac gtg agc acc cgc 960
Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg
305 310 315 320

ccc agc atc ggc agc aac gac atc atc acc agc ccc ttc tac ggc aac 1008
Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn
325 330 335

aag agc agc gag ccc gtg cag aac ctt gag ttc aac ggc gag aag gtg 1056
Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val
340 345 350

tac cgc gcc gtg gct aac acc aac ctg gcc gtg tgg ccc tct gca gtg 1104
Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val
355 360 365

tac agc ggc gtg acc aag gtg gag ttc agc cag tac aac gac cag acc 1152
Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr
370 375 380

gac gag gcc agc acc cag acc tac gac agc aag cgc aac gtg ggc gcc 1200
Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala
385 390 395 400

gtg agc tgg gac agc atc gac cag ctg ccc ccc gag acc acc gac gag 1248
Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu
405 410 415

ccc ctg gag aag ggc tac agc cac cag ctg aac tac gtg atg tgc ttc 1296
Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe
420 425 430

ctg atg cag ggc agc cgc ggc acc atc ccc gtg ctg acc tgg acc cac 1344
Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His
435 440 445

aag agc gtc gac ttc ttc aac atg atc gac agc aag aag atc acc cag 1392
Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln
450 455 460

ctg ccc ctg gtg aag gcc tac aag ctc cag agc ggc gcc agc gtg gtg 1440
Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val
465 470 475 480

gca ggc ccc cgc ttc acc ggc ggc gac atc atc cag tgc acc gag aac 1488
Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn
485 490 495

ggc agc gcc gcc acc atc tac gtg acc ccc gac gtg agc tac agc cag 1536
Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln
500 505 510

aag tac cgc gcc cgc atc cac tac gcc agc acc agc cag atc acc ttc 1584
Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe
515 520 525

acc ctg agc ctg gac ggg gcc ccc ttc aac caa tac tac ttc gac aag 1632
Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp Lys
530 535 540

acc atc aac aag ggc gac acc ctg acc tac aac agc ttc aac ctg gcc 1680
Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala
545 550 555 560

agc ttc agc acc cct ttc gag ctg agc ggc aac aac ctc cag atc ggc 1728
Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly
565 570 575

gtg acc ggc ctg agc gcc ggc gac aag gtg tac atc gac aag atc gag 1776
Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu
580 585 590

ttc atc ccc gtg aac tag atctgagct 1803
Phe Ile Pro Val Asn
595

<210> 4
<211> 597
<212> PRT
<213> Cry3A codificada por SEC ID NO:3.

<400> 4
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

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

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn
100 105 110

Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser
115 120 125

His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val
130 135 140

Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu

~~205~~
207

145 150 155 160

Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu
165 170 175

Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr
180 185 190

Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg
195 200 205

Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu
210 215 220

Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp
225 230 235 240

Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val
245 250 255

Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr
260 265 270

Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp
275 280 285

Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr
290 295 300

Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg
305 310 315 320

Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn
325 330 335

Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val
340 345 350

Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val
355 360 365

Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr
370 375 380

Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala
385 390 395 400

Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu
405 410 415

Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe
420 425 430

Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His
435 440 445

Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln
450 455 460

Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val
465 470 475 480

Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn
485 490 495

Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln
500 505 510

Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe
515 520 525

Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp Lys
530 535 540

Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala
545 550 555 560

Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly
565 570 575

Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu
580 585 590

Phe Ile Pro Val Asn
595

<210> 5
<211> 7208
<212> ADN
<213> Secuencia Artificial

<220>
<221> misc_carácter
<223> pCIB6850

<400> 5
gatccaccat gacggccgac aacaacaccg aggccctgga cagcagcacc accaaggacg 60
tgatccagaa gggcatcagc gtggtgggcg acctgctggg cgtggtgggc ttccccttcg 120
gcggcgccct ggtgagcttc tacaccaact tcctgaacac catctggccc agcgaggacc 180
cctggaaggc cttcatggag caggtggagg ccctgatgga ccagaagatc gccgactacg 240
ccaagaacaa ggcactggcc gagctacagg gcctccagaa caacgtggag gactatgtga 300
gcgccctgag cagctggcag aagaacccca tctcgagccg caacccccac agccagggcc 360
gcatccgcga gctgttcagc caggccgaga gccacttccg caacagcatg cccagcttcg 420
ccatcagcgg ctacgaggtg ctgttcctga ccacctacgc ccaggccgcc aacacccacc 480
tgttcctgct gaaggacgcc caaatctacg gagaggagtg gggctacgag aaggaggaca 540
tcgccgagtt ctacaagcgc cagctgaagc tgacccagga gtacaccgac cactgcgtga 600
agtggtacaa cgtgggtcta gacaagctcc gcggcagcag ctacgagagc tgggtgaact 660
tcaaccgcta ccgccgcgag atgaccctga ccgtgctgga cctgatcgcc ctgttccccc 720
tgtacgacgt gcgcctgtac cccaaggagg tgaagaccga gctgacccgc gacgtgctga 780
ccgaccccat cgtgggcgtg aacaacctgc gcggctacgg caccaccttc agcaacatcg 840
agaactacat ccgcaagccc cacctgttcg actacctgca ccgcatccag ttccacacgc 900
gtttccagcc cggctactac ggcaacgaca gcttcaacta ctggagcggc aactacgtga 960
gcacccgccc cagcatcggc agcaacgaca tcatcaccag ccccttctac ggcaacaaga 1020
gcagcgagcc cgtgcagaac cttgagttca acggcgagaa ggtgtaccgc gccgtggcta 1080
acaccaacct ggccgtgtgg ccctctgcag tgtacagcgg cgtgaccaag gtggagttca 1140
gccagtacaa cgaccagacc gacgaggcca gcacccagac ctacgacagc aagcgcaacg 1200
tgggcgccgt gagctgggac agcatcgacc agctgccccc cgagaccacc gacgagcccc 1260
tggagaaggg ctacagccac cagctgaact acgtgatgtg cttcctgatg cagggcagcc 1320
gcggcaccat ccccgtgctg acctggaccc acaagagcgt cgacttcttc aacatgatcg 1380

acagcaagaa gatcacccag ctgcccctgg tgaaggccta caagctccag agcggcgcca 1440
gcgtggtggc aggcccccgc ttcaccggcg gcgacatcat ccagtgcacc gagaacggca 1500
gcgccgccac catctacgtg acccccgacg tgagctacag ccagaagtac cgcgcccgca 1560
tccactacgc cagcaccagc cagatcacct tcaccctgag cctggacggg gcccccttca 1620
accaatacta cttcgacaag accatcaaca agggcgacac cctgacctac aacagcttca 1680
acctggccag cttcagcacc cctttcgagc tgagcggcaa caacctccag atcggcgtga 1740
ccggcctgag cgccggcgac aaggtgtaca tcgacaagat cgagttcatc cccgtgaact 1800
agatctgagc tcaagatctg ttgtacaaaa accagcaact cactgcactg cacttcactt 1860
cacttcactg tatgaataaa agtctggtgt ctggttcctg atcgatgact gactactcca 1920
ctttgtgcag aacttagtat gtatttgtat ttgtaaaata cttctatcaa taaaatttct 1980
aattcctaaa accaaaatcc agtgggtacc gaattcactg gccgtcgttt tacaacgtcg 2040
tgactgggaa aaccctggcg ttacccaact taatcgcctt gcagcacatc cccctttcgc 2100
cagctggcgt aatagcgaag aggcccgcac cgatcgccct tcccaacagt tgcgcagcct 2160
gaatggcgaa tggcgcctga tgcggtattt tctccttacg catctgtgcg gtatttcaca 2220
ccgcatatgg tgcactctca gtacaatctg ctctgatgcc gcatagttaa gccagccccg 2280
acacccgcca acacccgctg acgcgccctg acgggcttgt ctgctcccgg catccgctta 2340
cagacaagct gtgaccgtct ccgggagctg catgtgtcag aggttttcac cgtcatcacc 2400
gaaacgcgcg agacgaaagg gcctcgtgat acgcctattt ttataggtta atgtcatgat 2460
aataatggtt tcttagacgt caggtggcac ttttcgggga aatgtgcgcg gaacccctat 2520
ttgtttattt ttctaaatac attcaaatat gtatccgctc atgagacaat aaccctgata 2580
aatgcttcaa taatattgaa aaaggaagag tatgagtatt caacatttcc gtgtcgccct 2640
tattcccttt tttgcggcat tttgccttcc tgtttttgct cacccagaaa cgctggtgaa 2700
agtaaaagat gctgaagatc agttgggtgc acgagtgggt tacatcgaac tggatctcaa 2760
cagcggtaag atccttgaga gttttcgccc cgaagaacgt tttccaatga tgagcacttt 2820
taaagttctg ctatgtggcg cggtattatc ccgtattgac gccgggcaag agcaactcgg 2880
tcgccgcata cactattctc agaatgactt ggttgagtac tcaccagtca cagaaaagca 2940
tcttacggat ggcatgacag taagagaatt atgcagtgct gccataacca tgagtgataa 3000
cactgcggcc aacttacttc tgacaacgat cggaggaccg aaggagctaa ccgcttttt 3060
gcacaacatg ggggatcatg taactcgcct tgatcgttgg gaaccggagc tgaatgaagc 3120
cataccaaac gacgagcgtg acaccacgat gcctgtagca atggcaacaa cgttgcgcaa 3180
actattaact ggcgaactac ttactctagc ttcccggcaa caattaatag actggatgga 3240
ggcggataaa gttgcaggac cacttctgcg ctcggccctt ccggctggct ggtttattgc 3300
tgataaatct ggagccggtg agcgtgggtc tcgcggtatc attgcagcac tggggccaga 3360
tggtaagccc tcccgtatcg tagttatcta cacgacgggg agtcaggcaa ctatggatga 3420
acgaaataga cagatcgctg agataggtgc ctcactgatt aagcattggt aactgtcaga 3480
ccaagtttac tcatatatac tttagattga tttaaaactt catttttaat ttaaaaggat 3540
ctaggtgaag atccttttig ataatctcat gaccaaaatc ccttaacgtg agttttcgtt 3600
ccactgagcg tcagaccccg tagaaaagat caaaggatct tcttgagatc ctttttttct 3660
gcgcgtaatc tgctgcttgc aaacaaaaaa accaccgcta ccagcggtgg tttgtttgcc 3720
ggatcaagag ctaccaactc ttttccgaa ggtaactggc ttcagcagag cgcagatacc 3780

aaatactgtc cttctagtgt agccgtagtt aggccaccac ttcaagaact ctgtagcacc 3840
gcctacatac ctcgctctgc taatcctgtt accagtggct gctgccagtg gcgataagtc 3900
gtgtcttacc gggttggact caagacgata gttaccggat aaggcgcagc ggtcgggctg 3960
aacgggggggt tcgtgcacac agcccagctt ggagcgaacg acctacaccg aactgagata 4020
cctacagcgt gagctatgag aaagcgccac gcttcccgaa gggagaaagg cggacaggta 4080
tccggtaagc ggcagggtcg gaacaggaga gcgcacgagg gagcttccag ggggaaacgc 4140
ctggtatctt tatagtcctg tcgggtttcg ccacctctga cttgagcgtc gatttttgtg 4200
atgctcgtca ggggggcgga gcctatggaa aaacgccagc aacgcggcct ttttacggtt 4260
cctggccttt tgctggcctt ttgctcacat gttctttcct gcgttatccc ctgattctgt 4320
ggataaccgt attaccgcct ttgagtgagc tgataccgct cgccgcagcc gaacgaccga 4380
gcgcagcgag tcagtgagcg aggaagcgga agagcgccca atacgcaaac cgcctctccc 4440
cgcgcgttgg ccgattcatt aatgcagctg gcacgacagg tttcccgact ggaaagcggg 4500
cagtgagcgc aacgcaatta atgtgagtta gctcactcat taggcacccc aggctttaca 4560
ctttatgctt ccggctcgta tgttgtgtgg aattgtgagc ggataacaat ttcacacagg 4620
aaacagctat gaccatgatt acgccaagct tgcacatgac aacaattgta agaggatgga 4680
gaccacaacg atccaacaat acttctgcga cgggctgtga agtatagaga agttaaacgc 4740
ccaaaagcca ttgtgtttgg aatttttagt tattctattt ttcatgatgt atcttcctct 4800
aacatgcctt aatttgcaaa tttggtataa ctactgattg aaaatatatg tatgtaaaaa 4860
aatactaagc atatttgtga agctaaacat gatgttaitt aagaaaatat gttgttaaca 4920
gaataagatt aatatcgaaa tggaaacatc tgtaaattag aatcatctta caagctaaga 4980
gatgttcacg ctttgagaaa cttcttcaga tcatgaccgt agaagtagct ctccaagact 5040
caacgaaggc tgctgcaatt ccacaaatgc atgacatgca tccttgtaac cgtcgtcgcc 5100
gctataaaca cggataactc aattccctgc tccatcaatt tagaaatgag caagcaagca 5160
cccgatcgct caccccatat gcaccaatct gactcccaag tctctgtttc gcattagtac 5220
cgccagcact ccacctatag ctaccaattg agacctttcc agcctaagca gatcgattga 5280
tcgttagagt caaagagttg gtggtacggg tactttaact accatggaat gatggggcgt 5340
gatgtagagc ggaaagcgcc tccctacgcg gaacaacacc ctcgccatgc cgctcgacta 5400
cagcctcctc ctcgtcggcc gcccacaacg agggagcccg tggtcgcagc caccgaccag 5460
catgtctctg tgtcctcgtc cgacctcgac atgtcatggc aaacagtcgg acgccagcac 5520
cagactgacg acatgagtct ctgaagagcc cgccacctag aaagatccga gccctgctgc 5580
tggtagtggt aaccattttc gtcgcgctga cgcggagagc gagaggccag aaatttatag 5640
cgactgacgc tgtggcaggc acgctatcgg aggttacgac gtggcgggtc actcgacgcg 5700
gagttcacag gtcctatcct tgcatcgctc gggccggagt ttacgggact tatccttacg 5760
acgtgctcta aggttgcgat aacgggcgga ggaaggcgtg tggcgtgcgg agacggttta 5820
tacacgtagt gtgcgggagt gtgtttcgta gacgcgggaa agcacgacga cttacgaagg 5880
ttagtggagg aggaggacac actaaaatca ggacgcaaga aactcttcta ttatagtagt 5940
agagaagaga ttataggagt gtgggttgat tctaaagaaa atcgacgcag gacaaccgtc 6000
aaaacgggtg ctttaatata gtagatatat atatatagag agagagagaa agtacaaagg 6060
atgcatttgt gtctgcatat gatcggagta ttactaacgg ccgtcgtaag aaggtccatc 6120
atgcgtggag cgagcccatt tggttggttg tcaggccgca gttaaggcct ccatatatga 6180

ttgtcgtcgg gcccataaca gcatctcctc caccagttta ttgtaagaat aaattaagta 6240
gagatatttg tcgtcgggca gaagaaactt ggacaagaag aagaagcaag ctaggccaat 6300
ttcttgccgg caagaggaag atagtggcct ctagtttata tatcggcgtg atgatgatgc 6360
tcctagctag aaatgagaga agaaaaacgg acgcgtgttt ggtgtgtgtc aatggcgtcc 6420
atccttccat cagatcagaa cgatgaaaaa gtcaagcacg gcatgcatag tatatgtata 6480
gcttgtttta gtgtggcttt gctgagacga atgaaagcaa cggcgggcat atttttcagt 6540
ggctgtagct ttcaggctga aagagacgtg gcatgcaata attcagggaa ttcgtcagcc 6600
aattgaggta gctagtcaac ttgtacattg gtgcgagcaa ttttccgcac tcaggagggc 6660
tagtttgaga gtccaaaaac tataggagat taaagaggct aaaatcctct ccttatttaa 6720
ttttaaataa gtagtgtatt tgtattttaa ctcctccaac ccttccgatt ttatggctct 6780
caaactagca ttcagtctaa tgcatgcatg cttggctaga ggtcgtatgg ggttgttaat 6840
agcatagcta gctacaagtt aaccgggtct tttatattta ataaggacag gcaaagtatt 6900
acttacaaat aaagaataaa gctaggacga actcgtggat tattactaaa tcgaaatgga 6960
cgtaatattc caggcaagaa taattgttcg atcaggagac aagtggggca ttggaccggt 7020
tcttgcaagc aagagcctat ggcgtggtga cacggcgcgt tgcccataca tcatgcctcc 7080
atcgatgatc catcctcact tgctataaaa agaggtgtcc atggtgctca agctcagcca 7140
agcaaataag acgacttgtt tcattgattc ttcaagagat cgagcttctt ttgcaccaca 7200
aggtcgag 7208

<210> 6
<211> 1801
<212> ADN
<213> Secuencia Artificial

<220>
<221> CDS
<222> (1)..(1791)
<223> Secuencia codificadora de cry3A054 de maíz optimizado.

<220>
<221> misc_carácter
<222> (322)..(333)
<223> Secuencia codificadora del sito de reconocimiento de catepsina G

<400> 6
atg acg gcc gac aac aac acc gag gcc ctg gac agc agc acc acc aag 48
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

gac gtg atc cag aag ggc atc agc gtg gtg ggc gac ctg ctg ggc gtg 96
Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

gtg ggc ttc ccc ttc ggc ggc gcc ctg gtg agc ttc tac acc aac ttc 144
Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

ctg aac acc atc tgg ccc agc gag gac ccc tgg aag gcc ttc atg gag 192
Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

cag gtg gag gcc ctg atg gac cag aag atc gcc gac tac gcc aag aac 240
Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65 70 75 80

aag gca ctg gcc gag cta cag ggc ctc cag aac aac gtg gag gac tat 288
Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

gtg agc gcc ctg agc agc tgg cag aag aac ccc gct gca ccg ttc ccc 336
Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Pro
100 105 110

cac agc cag ggc cgc atc cgc gag ctg ttc agc cag gcc gag agc cac 384
His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser His
115 120 125

ttc cgc aac agc atg ccc agc ttc gcc atc agc ggc tac gag gtg ctg 432
Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val Leu
130 135 140

ttc ctg acc acc tac gcc cag gcc gcc aac acc cac ctg ttc ctg ctg 480
Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu Leu
145 150 155 160

aag gac gcc caa atc tac gga gag gag tgg ggc tac gag aag gag gac 528
Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu Asp
165 170 175

atc gcc gag ttc tac aag cgc cag ctg aag ctg acc cag gag tac acc 576
Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr Thr
180 185 190

gac cac tgc gtg aag tgg tac aac gtg ggt cta gac aag ctc cgc ggc 624
Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg Gly
195 200 205

agc agc tac gag agc tgg gtg aac ttc aac cgc tac cgc cgc gag atg 672
Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu Met
210 215 220

acc ctg acc gtg ctg gac ctg atc gcc ctg ttc ccc ctg tac gac gtg 720
Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp Val
225 230 235 240

cgc ctg tac ccc aag gag gtg aag acc gag ctg acc cgc gac gtg ctg 768
Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val Leu
245 250 255

acc gac ccc atc gtg ggc gtg aac aac ctg cgc ggc tac ggc acc acc 816
Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr Thr
260 265 270

ttc agc aac atc gag aac tac atc cgc aag ccc cac ctg ttc gac tac 864
Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp Tyr
275 280 285

ctg cac cgc atc cag ttc cac acg cgt ttc cag ccc ggc tac tac ggc 912
Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr Gly
290 295 300

aac gac agc ttc aac tac tgg agc ggc aac tac gtg agc acc cgc ccc 960
Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg Pro
305 310 315 320

agc atc ggc agc aac gac atc atc acc agc ccc ttc tac ggc aac aag 1008
Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn Lys
325 330 335

agc agc gag ccc gtg cag aac ctt gag ttc aac ggc gag aag gtg tac 1056
Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val Tyr
340 345 350

cgc gcc gtg gct aac acc aac ctg gcc gtg tgg ccc tct gca gtg tac 1104
Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val Tyr
355 360 365

agc ggc gtg acc aag gtg gag ttc agc cag tac aac gac cag acc gac 1152
Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr Asp
370 375 380

gag gcc agc acc cag acc tac gac agc aag cgc aac gtg ggc gcc gtg 1200
Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala Val
385 390 395 400

agc tgg gac agc atc gac cag ctg ccc ccc gag acc acc gac gag ccc 1248
Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu Pro
405 410 415

ctg gag aag ggc tac agc cac cag ctg aac tac gtg atg tgc ttc ctg 1296
Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe Leu
420 425 430

atg cag ggc agc cgc ggc acc atc ccc gtg ctg acc tgg acc cac aag 1344
Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His Lys
435 440 445

agc gtc gac ttc ttc aac atg atc gac agc aag aag atc acc cag ctg 1392
Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln Leu
450 455 460

ccc ctg gtg aag gcc tac aag ctc cag agc ggc gcc agc gtg gtg gca 1440
Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val Ala
465 470 475 480

ggc ccc cgc ttc acc ggc ggc gac atc atc cag tgc acc gag aac ggc 1488
Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn Gly
485 490 495

agc gcc gcc acc atc tac gtg acc ccc gac gtg agc tac agc cag aag 1536
Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln Lys
500 505 510

tac cgc gcc cgc atc cac tac gcc agc acc agc cag atc acc ttc acc 1584
Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe Thr
515 520 525

ctg agc ctg gac ggg gcc ccc ttc aac caa tac tac ttc gac aag acc 1632
Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp Lys Thr
530 535 540

atc aac aag ggc gac acc ctg acc tac aac agc ttc aac ctg gcc agc 1680
Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala Ser
545 550 555 560

ttc agc acc cct ttc gag ctg agc ggc aac aac ctc cag atc ggc gtg 1728
Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly Val
565 570 575

acc ggc ctg agc gcc ggc gac aag gtg tac atc gac aag atc gag ttc 1776
Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu Phe
580 585 590

atc ccc gtg aac tag atctgagctc 1801
Ile Pro Val Asn
595

<210> 7
<211> 596
<212> PRT
<213> Secuencia Artificial

<220>
<221> misc_carácter
<222> (322)..(333)
<223> Secuencia codificadora del sitio de reconocimiento de catepsina G

<400> 7

Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

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

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Pro
100 105 110

His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser His
115 120 125

Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val Leu
130 135 140

Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu Leu
145 150 155 160

Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu Asp
165 170 175

Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr Thr
180 185 190

Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg Gly
195 200 205

Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu Met

210 215 220

Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp Val
225 230 235 240

Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val Leu
245 250 255

Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr Thr
260 265 270

Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp Tyr
275 280 285

Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr Gly
290 295 300

Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg Pro
305 310 315 320

Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn Lys
325 330 335

Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val Tyr
340 345 350

Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val Tyr
355 360 365

Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr Asp
370 375 380

Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala Val
385 390 395 400

Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu Pro
405 410 415

Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe Leu
420 425 430

Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His Lys
435 440 445

Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln Leu
450 455 460

Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val Ala
465 470 475 480

Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn Gly
485 490 495

Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln Lys
500 505 510

Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe Thr
515 520 525

Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp Lys Thr
530 535 540

Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala Ser
545 550 555 560

Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly Val
565 570 575

Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu Phe
580 585 590

Ile Pro Val Asn
595

<210> 8
<211> 1807
<212> ADN
<213> Secuencia Artificial
<220>

220 218

<221> CDS
<222> (1)..(1806)
<223> Secuencia codificadora de cry3A055 modificada de maíz optimizado
<220>
<221> misc_carácter
<222> (322)..(333)
<223> Secuencia codificadora del sitio de reconocimiento de Catepsina G.

<400> 8
atg acg gcc gac aac aac acc gag gcc ctg gac agc agc acc acc aag 48
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

gac gtg atc cag aag ggc atc agc gtg gtg ggc gac ctg ctg ggc gtg 96
Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

gtg ggc ttc ccc ttc ggc ggc gcc ctg gtg agc ttc tac acc aac ttc 144
Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

ctg aac acc atc tgg ccc agc gag gac ccc tgg aag gcc ttc atg gag 192
Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

cag gtg gag gcc ctg atg gac cag aag atc gcc gac tac gcc aag aac 240
Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65 70 75 80

aag gca ctg gcc gag cta cag ggc ctc cag aac aac gtg gag gac tat 288
Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

gtg agc gcc ctg agc agc tgg cag aag aac ccc gct gca ccg ttc cgc 336
Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Arg
100 105 110

aac ccc cac agc cag ggc cgc atc cgc gag ctg ttc agc cag gcc gag 384
Asn Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu
115 120 125

agc cac ttc cgc aac agc atg ccc agc ttc gcc atc agc ggc tac gag 432
Ser His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu
130 135 140

gtg ctg ttc ctg acc acc tac gcc cag gcc gcc aac acc cac ctg ttc 480
Val Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe
145 150 155 160

ctg ctg aag gac gcc caa atc tac gga gag gag tgg ggc tac gag aag 528
Leu Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys
165 170 175

gag gac atc gcc gag ttc tac aag cgc cag ctg aag ctg acc cag gag 576
Glu Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu
180 185 190

tac acc gac cac tgc gtg aag tgg tac aac gtg ggt cta gac aag ctc 624
Tyr Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu
195 200 205

cgc ggc agc agc tac gag agc tgg gtg aac ttc aac cgc tac cgc cgc 672
Arg Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg
210 215 220

gag atg acc ctg acc gtg ctg gac ctg atc gcc ctg ttc ccc ctg tac 720
Glu Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr
225 230 235 240

gac gtg cgc ctg tac ccc aag gag gtg aag acc gag ctg acc cgc gac 768
Asp Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp
245 250 255

gtg ctg acc gac ccc atc gtg ggc gtg aac aac ctg cgc ggc tac ggc 816
Val Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly
260 265 270

acc acc ttc agc aac atc gag aac tac atc cgc aag ccc cac ctg ttc 864
Thr Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe
275 280 285

gac tac ctg cac cgc atc cag ttc cac acg cgt ttc cag ccc ggc tac 912
Asp Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr
290 295 300

tac ggc aac gac agc ttc aac tac tgg agc ggc aac tac gtg agc acc 960
Tyr Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr
305 310 315 320

cgc ccc agc atc ggc agc aac gac atc atc acc agc ccc ttc tac ggc 1008
Arg Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly
325 330 335

aac aag agc agc gag ccc gtg cag aac ctt gag ttc aac ggc gag aag 1056
Asn Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys
340 345 350

gtg tac cgc gcc gtg gct aac acc aac ctg gcc gtg tgg ccc tct gca 1104
Val Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala
355 360 365

gtg tac agc ggc gtg acc aag gtg gag ttc agc cag tac aac gac cag 1152
Val Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln
370 375 380

acc gac gag gcc agc acc cag acc tac gac agc aag cgc aac gtg ggc 1200
Thr Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly
385 390 395 400

gcc gtg agc tgg gac agc atc gac cag ctg ccc ccc gag acc acc gac 1248
Ala Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp
405 410 415

gag ccc ctg gag aag ggc tac agc cac cag ctg aac tac gtg atg tgc 1296
Glu Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys
420 425 430

ttc ctg atg cag ggc agc cgc ggc acc atc ccc gtg ctg acc tgg acc 1344
Phe Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr
435 440 445

cac aag agc gtc gac ttc ttc aac atg atc gac agc aag aag atc acc 1392
His Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr
450 455 460

cag ctg ccc ctg gtg aag gcc tac aag ctc cag agc ggc gcc agc gtg 1440
Gln Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val
465 470 475 480

gtg gca ggc ccc cgc ttc acc ggc ggc gac atc atc cag tgc acc gag 1488
Val Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu
485 490 495

aac ggc agc gcc gcc acc atc tac gtg acc ccc gac gtg agc tac agc 1536
Asn Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser
500 505 510

cag aag tac cgc gcc cgc atc cac tac gcc agc acc agc cag atc acc 1584
Gln Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr
515 520 525

ttc acc ctg agc ctg gac ggg gcc ccc ttc aac caa tac tac ttc gac 1632
Phe Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp
530 535 540

aag acc atc aac aag ggc gac acc ctg acc tac aac agc ttc aac ctg 1680
Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu
545 550 555 560

gcc agc ttc agc acc cct ttc gag ctg agc ggc aac aac ctc cag atc 1728
Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile
565 570 575

ggc gtg acc ggc ctg agc gcc ggc gac aag gtg tac atc gac aag atc 1776
Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile
580 585 590

gag ttc atc ccc gtg aac tag atc tga gct c 1807
Glu Phe Ile Pro Val Asn Ile Ala
595 600

<210> 9
<211> 598
<212> PRT
<213> Secuencia Artificial

<220>
<221> misc_carácter
<222> (322)..(333)
<223> Secuencia codificadora del sitio de reconocimiento de Catepsina G.

<400> 9
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

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

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Arg
100 105 110

Asn Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu
115 120 125

Ser His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu
130 135 140

Val Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe
145 150 155 160

Leu Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys
165 170 175

Glu Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu
180 185 190

Tyr Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu
195 200 205

Arg Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg
210 215 220

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

Asp Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp
245 250 255

Val Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly
260 265 270

Thr Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe
275 280 285

Asp Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr
290 295 300

Tyr Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr
305 310 315 320

Arg Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly
325 330 335

Asn Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys
340 345 350

Val Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala

355 360 365

Val Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln
370 375 380

Thr Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly
385 390 395 400

Ala Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp
405 410 415

Glu Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys
420 425 430

Phe Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr
435 440 445

His Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr
450 455 460

Gln Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val
465 470 475 480

Val Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu
485 490 495

Asn Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser
500 505 510

Gln Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr
515 520 525

Phe Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp
530 535 540

Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu
545 550 555 560

Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile
565 570 575

Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile
580 585 590

Glu Phe Ile Pro Val Asn
595

<210> 10
<211> 1818
<212> ADN
<213> Secuencia Artificial

<220>
<221> CDS
<222> (1)..(1818)
<223> Secuencia codificadora cry3A085 modificada optimizada de maíz.

<220>
<221> misc_carácter
<222> (346)..(357)
<223> Secuencia Codificadora del sitio de reconocimiento G.
<400> 10
atg aac tac aag gag ttc ctc cgc atg acc gcc gac aac aac acc gag 48
Met Asn Tyr Lys Glu Phe Leu Arg Met Thr Ala Asp Asn Asn Thr Glu
1 5 10 15

gcc ctg gac agc agc acc acc aag gac gtg atc cag aag ggc atc agc 96
Ala Leu Asp Ser Ser Thr Thr Lys Asp Val Ile Gln Lys Gly Ile Ser
20 25 30

gtg gtg ggc gac ctg ctg ggc gtg gtg ggc ttc ccc ttc ggc ggc gcc 144
Val Val Gly Asp Leu Leu Gly Val Val Gly Phe Pro Phe Gly Gly Ala
35 40 45

ctg gtg agc ttc tac acc aac ttc ctg aac acc atc tgg ccc agc gag 192
Leu Val Ser Phe Tyr Thr Asn Phe Leu Asn Thr Ile Trp Pro Ser Glu
50 55 60

gac ccc tgg aag gcc ttc atg gag cag gtg gag gcc ctg atg gac cag 240

Asp Pro Trp Lys Ala Phe Met Glu Gln Val Glu Ala Leu Met Asp Gln
65 70 75 80

aag atc gcc gac tac gcc aag aac aag gca ctg gcc gag cta cag ggc 288
Lys Ile Ala Asp Tyr Ala Lys Asn Lys Ala Leu Ala Glu Leu Gln Gly
85 90 95

ctc cag aac aac gtg gag gac tat gtg agc gcc ctg agc agc tgg cag 336
Leu Gln Asn Asn Val Glu Asp Tyr Val Ser Ala Leu Ser Ser Trp Gln
100 105 110

aag aac ccc gct gca ccg ttc cgc aac ccc cac agc cag ggc cgc atc 384
Lys Asn Pro Ala Ala Pro Phe Arg Asn Pro His Ser Gln Gly Arg Ile
115 120 125

cgc gag ctg ttc agc cag gcc gag agc cac ttc cgc aac agc atg ccc 432
Arg Glu Leu Phe Ser Gln Ala Glu Ser His Phe Arg Asn Ser Met Pro
130 135 140

agc ttc gcc atc agc ggc tac gag gtg ctg ttc ctg acc acc tac gcc 480
Ser Phe Ala Ile Ser Gly Tyr Glu Val Leu Phe Leu Thr Thr Tyr Ala
145 150 155 160

cag gcc gcc aac acc cac ctg ttc ctg ctg aag gac gcc caa atc tac 528
Gln Ala Ala Asn Thr His Leu Phe Leu Leu Lys Asp Ala Gln Ile Tyr
165 170 175

gga gag gag tgg ggc tac gag aag gag gac atc gcc gag ttc tac aag 576
Gly Glu Glu Trp Gly Tyr Glu Lys Glu Asp Ile Ala Glu Phe Tyr Lys
180 185 190

cgc cag ctg aag ctg acc cag gag tac acc gac cac tgc gtg aag tgg 624
Arg Gln Leu Lys Leu Thr Gln Glu Tyr Thr Asp His Cys Val Lys Trp
195 200 205

tac aac gtg ggt cta gac aag ctc cgc ggc agc agc tac gag agc tgg 672
Tyr Asn Val Gly Leu Asp Lys Leu Arg Gly Ser Ser Tyr Glu Ser Trp
210 215 220

gtg aac ttc aac cgc tac cgc cgc gag atg acc ctg acc gtg ctg gac 720

Val Asn Phe Asn Arg Tyr Arg Arg Glu Met Thr Leu Thr Val Leu Asp
225 230 235 240

ctg atc gcc ctg ttc ccc ctg tac gac gtg cgc ctg tac ccc aag gag 768
Leu Ile Ala Leu Phe Pro Leu Tyr Asp Val Arg Leu Tyr Pro Lys Glu
245 250 255

gtg aag acc gag ctg acc cgc gac gtg ctg acc gac ccc atc gtg ggc 816
Val Lys Thr Glu Leu Thr Arg Asp Val Leu Thr Asp Pro Ile Val Gly
260 265 270

gtg aac aac ctg cgc ggc tac ggc acc acc ttc agc aac atc gag aac 864
Val Asn Asn Leu Arg Gly Tyr Gly Thr Thr Phe Ser Asn Ile Glu Asn
275 280 285

tac atc cgc aag ccc cac ctg ttc gac tac ctg cac cgc atc cag ttc 912
Tyr Ile Arg Lys Pro His Leu Phe Asp Tyr Leu His Arg Ile Gln Phe
290 295 300

cac acg cgt ttc cag ccc ggc tac tac ggc aac gac agc ttc aac tac 960
His Thr Arg Phe Gln Pro Gly Tyr Tyr Gly Asn Asp Ser Phe Asn Tyr
305 310 315 320

tgg agc ggc aac tac gtg agc acc cgc ccc agc atc ggc agc aac gac 1008
Trp Ser Gly Asn Tyr Val Ser Thr Arg Pro Ser Ile Gly Ser Asn Asp
325 330 335

atc atc acc agc ccc ttc tac ggc aac aag agc agc gag ccc gtg cag 1056
Ile Ile Thr Ser Pro Phe Tyr Gly Asn Lys Ser Ser Glu Pro Val Gln
340 345 350

aac ctt gag ttc aac ggc gag aag gtg tac cgc gcc gtg gct aac acc 1104
Asn Leu Glu Phe Asn Gly Glu Lys Val Tyr Arg Ala Val Ala Asn Thr
355 360 365

aac ctg gcc gtg tgg ccc tct gca gtg tac agc ggc gtg acc aag gtg 1152
Asn Leu Ala Val Trp Pro Ser Ala Val Tyr Ser Gly Val Thr Lys Val
370 375 380

gag ttc agc cag tac aac gac cag acc gac gag gcc agc acc cag acc 1200

Glu Phe Ser Gln Tyr Asn Asp Gln Thr Asp Glu Ala Ser Thr Gln Thr
385 390 395 400

tac gac agc aag cgc aac gtg ggc gcc gtg agc tgg gac agc atc gac 1248
Tyr Asp Ser Lys Arg Asn Val Gly Ala Val Ser Trp Asp Ser Ile Asp
405 410 415

cag ctg ccc ccc gag acc acc gac gag ccc ctg gag aag ggc tac agc 1296
Gln Leu Pro Pro Glu Thr Thr Asp Glu Pro Leu Glu Lys Gly Tyr Ser
420 425 430

cac cag ctg aac tac gtg atg tgc ttc ctg atg cag ggc agc cgc ggc 1344
His Gln Leu Asn Tyr Val Met Cys Phe Leu Met Gln Gly Ser Arg Gly
435 440 445

acc atc ccc gtg ctg acc tgg acc cac aag agc gtc gac ttc ttc aac 1392
Thr Ile Pro Val Leu Thr Trp Thr His Lys Ser Val Asp Phe Phe Asn
450 455 460

atg atc gac agc aag aag atc acc cag ctg ccc ctg gtg aag gcc tac 1440
Met Ile Asp Ser Lys Lys Ile Thr Gln Leu Pro Leu Val Lys Ala Tyr
465 470 475 480

aag ctc cag agc ggc gcc agc gtg gtg gca ggc ccc cgc ttc acc ggc 1488
Lys Leu Gln Ser Gly Ala Ser Val Val Ala Gly Pro Arg Phe Thr Gly
485 490 495

ggc gac atc atc cag tgc acc gag aac ggc agc gcc gcc acc atc tac 1536
Gly Asp Ile Ile Gln Cys Thr Glu Asn Gly Ser Ala Ala Thr Ile Tyr
500 505 510

gtg acc ccc gac gtg agc tac agc cag aag tac cgc gcc cgc atc cac 1584
Val Thr Pro Asp Val Ser Tyr Ser Gln Lys Tyr Arg Ala Arg Ile His
515 520 525

tac gcc agc acc agc cag atc acc ttc acc ctg agc ctg gac ggg gcc 1632
Tyr Ala Ser Thr Ser Gln Ile Thr Phe Thr Leu Ser Leu Asp Gly Ala
530 535 540

ccc ttc aac caa tac tac ttc gac aag acc atc aac aag ggc gac acc 1680

Pro Phe Asn Gln Tyr Tyr Phe Asp Lys Thr Ile Asn Lys Gly Asp Thr
545 550 555 560

ctg acc tac aac agc ttc aac ctg gcc agc ttc agc acc cct ttc gag 1728
Leu Thr Tyr Asn Ser Phe Asn Leu Ala Ser Phe Ser Thr Pro Phe Glu
565 570 575

ctg agc ggc aac aac ctc cag atc ggc gtg acc ggc ctg agc gcc ggc 1776
Leu Ser Gly Asn Asn Leu Gln Ile Gly Val Thr Gly Leu Ser Ala Gly
580 585 590

gac aag gtg tac atc gac aag atc gag ttc atc ccc gtg aac 1818
Asp Lys Val Tyr Ile Asp Lys Ile Glu Phe Ile Pro Val Asn
595 600 605

<210> 11
<211> 606
<212> PRT
<213> Secuencia Artificial

<220>
<221> misc_carácter
<222> (346)..(357)
<223> Secuencia codificadora del sitio de reconocimiento de Catepsina G.

<400> 11
Met Asn Tyr Lys Glu Phe Leu Arg Met Thr Ala Asp Asn Asn Thr Glu
1 5 10 15

Ala Leu Asp Ser Ser Thr Thr Lys Asp Val Ile Gln Lys Gly Ile Ser
20 25 30

Val Val Gly Asp Leu Leu Gly Val Val Gly Phe Pro Phe Gly Gly Ala
35 40 45

Leu Val Ser Phe Tyr Thr Asn Phe Leu Asn Thr Ile Trp Pro Ser Glu
50 55 60

Asp Pro Trp Lys Ala Phe Met Glu Gln Val Glu Ala Leu Met Asp Gln

65 70 75 80

Lys Ile Ala Asp Tyr Ala Lys Asn Lys Ala Leu Ala Glu Leu Gln Gly
85 90 95

Leu Gln Asn Asn Val Glu Asp Tyr Val Ser Ala Leu Ser Ser Trp Gln
100 105 110

Lys Asn Pro Ala Ala Pro Phe Arg Asn Pro His Ser Gln Gly Arg Ile
115 120 125

Arg Glu Leu Phe Ser Gln Ala Glu Ser His Phe Arg Asn Ser Met Pro
130 135 140

Ser Phe Ala Ile Ser Gly Tyr Glu Val Leu Phe Leu Thr Thr Tyr Ala
145 150 155 160

Gln Ala Ala Asn Thr His Leu Phe Leu Leu Lys Asp Ala Gln Ile Tyr
165 170 175

Gly Glu Glu Trp Gly Tyr Glu Lys Glu Asp Ile Ala Glu Phe Tyr Lys
180 185 190

Arg Gln Leu Lys Leu Thr Gln Glu Tyr Thr Asp His Cys Val Lys Trp
195 200 205

Tyr Asn Val Gly Leu Asp Lys Leu Arg Gly Ser Ser Tyr Glu Ser Trp
210 215 220

Val Asn Phe Asn Arg Tyr Arg Arg Glu Met Thr Leu Thr Val Leu Asp
225 230 235 240

Leu Ile Ala Leu Phe Pro Leu Tyr Asp Val Arg Leu Tyr Pro Lys Glu
245 250 255

Val Lys Thr Glu Leu Thr Arg Asp Val Leu Thr Asp Pro Ile Val Gly
260 265 270

Val Asn Asn Leu Arg Gly Tyr Gly Thr Thr Phe Ser Asn Ile Glu Asn
275 280 285

Tyr Ile Arg Lys Pro His Leu Phe Asp Tyr Leu His Arg Ile Gln Phe
290 295 300

His Thr Arg Phe Gln Pro Gly Tyr Tyr Gly Asn Asp Ser Phe Asn Tyr
305 310 315 320

Trp Ser Gly Asn Tyr Val Ser Thr Arg Pro Ser Ile Gly Ser Asn Asp
325 330 335

Ile Ile Thr Ser Pro Phe Tyr Gly Asn Lys Ser Ser Glu Pro Val Gln
340 345 350

Asn Leu Glu Phe Asn Gly Glu Lys Val Tyr Arg Ala Val Ala Asn Thr
355 360 365

Asn Leu Ala Val Trp Pro Ser Ala Val Tyr Ser Gly Val Thr Lys Val
370 375 380

Glu Phe Ser Gln Tyr Asn Asp Gln Thr Asp Glu Ala Ser Thr Gln Thr
385 390 395 400

Tyr Asp Ser Lys Arg Asn Val Gly Ala Val Ser Trp Asp Ser Ile Asp
405 410 415

Gln Leu Pro Pro Glu Thr Thr Asp Glu Pro Leu Glu Lys Gly Tyr Ser
420 425 430

His Gln Leu Asn Tyr Val Met Cys Phe Leu Met Gln Gly Ser Arg Gly
435 440 445

Thr Ile Pro Val Leu Thr Trp Thr His Lys Ser Val Asp Phe Phe Asn
450 455 460

Met Ile Asp Ser Lys Lys Ile Thr Gln Leu Pro Leu Val Lys Ala Tyr
465 470 475 480

Lys Leu Gln Ser Gly Ala Ser Val Val Ala Gly Pro Arg Phe Thr Gly
485 490 495

Gly Asp Ile Ile Gln Cys Thr Glu Asn Gly Ser Ala Ala Thr Ile Tyr
500 505 510

Val Thr Pro Asp Val Ser Tyr Ser Gln Lys Tyr Arg Ala Arg Ile His
515 520 525

Tyr Ala Ser Thr Ser Gln Ile Thr Phe Thr Leu Ser Leu Asp Gly Ala
530 535 540

Pro Phe Asn Gln Tyr Tyr Phe Asp Lys Thr Ile Asn Lys Gly Asp Thr
545 550 555 560

Leu Thr Tyr Asn Ser Phe Asn Leu Ala Ser Phe Ser Thr Pro Phe Glu
565 570 575

Leu Ser Gly Asn Asn Leu Gln Ile Gly Val Thr Gly Leu Ser Ala Gly
580 585 590

Asp Lys Val Tyr Ile Asp Lys Ile Glu Phe Ile Pro Val Asn
595 600 605

<210> 12
<211> 1794
<212> ADN
<213> Secuencia Artificial

<220>
<221> CDS
<222> (1)..(1794)
<223> Secuencia codificadora de cry3A082 modificado optimizado.

<220>
<221> misc_carácter
<222> (1609)..(1620)
<223> Secuencia codificadora del sitio de reconocimiento de Catepsina G

<400> 12
atg acg gcc gac aac aac acc gag gcc ctg gac agc agc acc acc aag 48
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys

1 5 10 15

gac gtg atc cag aag ggc atc agc gtg gtg ggc gac ctg ctg ggc gtg 96
Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

gtg ggc ttc ccc ttc ggc ggc gcc ctg gtg agc ttc tac acc aac ttc 144
Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

ctg aac acc atc tgg ccc agc gag gac ccc tgg aag gcc ttc atg gag 192
Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

cag gtg gag gcc ctg atg gac cag aag atc gcc gac tac gcc aag aac 240
Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65 70 75 80

aag gca ctg gcc gag cta cag ggc ctc cag aac aac gtg gag gac tat 288
Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

gtg agc gcc ctg agc agc tgg cag aag aac ccc gtc tcg agc cgc aac 336
Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn
100 105 110

ccc cac agc cag ggc cgc atc cgc gag ctg ttc agc cag gcc gag agc 384
Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser
115 120 125

cac ttc cgc aac agc atg ccc agc ttc gcc atc agc ggc tac gag gtg 432
His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val
130 135 140

ctg ttc ctg acc acc tac gcc cag gcc gcc aac acc cac ctg ttc ctg 480
Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu
145 150 155 160

ctg aag gac gcc caa atc tac gga gag gag tgg ggc tac gag aag gag 528
Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu

165 170 175

gac atc gcc gag ttc tac aag cgc cag ctg aag ctg acc cag gag tac 576
Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr
180 185 190

acc gac cac tgc gtg aag tgg tac aac gtg ggt cta gac aag ctc cgc 624
Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg
195 200 205

ggc agc agc tac gag agc tgg gtg aac ttc aac cgc tac cgc cgc gag 672
Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu
210 215 220

atg acc ctg acc gtg ctg gac ctg atc gcc ctg ttc ccc ctg tac gac 720
Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp
225 230 235 240

gtg cgc ctg tac ccc aag gag gtg aag acc gag ctg acc cgc gac gtg 768
Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val
245 250 255

ctg acc gac ccc atc gtg ggc gtg aac aac ctg cgc ggc tac ggc acc 816
Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr
260 265 270

acc ttc agc aac atc gag aac tac atc cgc aag ccc cac ctg ttc gac 864
Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp
275 280 285

tac ctg cac cgc atc cag ttc cac acg cgt ttc cag ccc ggc tac tac 912
Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr
290 295 300

ggc aac gac agc ttc aac tac tgg agc ggc aac tac gtg agc acc cgc 960
Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg
305 310 315 320

ccc agc atc ggc agc aac gac atc atc acc agc ccc ttc tac ggc aac 1008
Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn

325 330 335

aag agc agc gag ccc gtg cag aac ctt gag ttc aac ggc gag aag gtg 1056
Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val
340 345 350

tac cgc gcc gtg gct aac acc aac ctg gcc gtg tgg ccc tct gca gtg 1104
Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val
355 360 365

tac agc ggc gtg acc aag gtg gag ttc agc cag tac aac gac cag acc 1152
Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr
370 375 380

gac gag gcc agc acc cag acc tac gac agc aag cgc aac gtg ggc gcc 1200
Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala
385 390 395 400

gtg agc tgg gac agc atc gac cag ctg ccc ccc gag acc acc gac gag 1248
Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu
405 410 415

ccc ctg gag aag ggc tac agc cac cag ctg aac tac gtg atg tgc ttc 1296
Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe
420 425 430

ctg atg cag ggc agc cgc ggc acc atc ccc gtg ctg acc tgg acc cac 1344
Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His
435 440 445

aag agc gtc gac ttc ttc aac atg atc gac agc aag aag atc acc cag 1392
Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln
450 455 460

ctg ccc ctg gtg aag gcc tac aag ctc cag agc ggc gcc agc gtg gtg 1440
Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val
465 470 475 480

gca ggc ccc cgc ttc acc ggc ggc gac atc atc cag tgc acc gag aac 1488
Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn

485 490 495

ggc agc gcc gcc acc atc tac gtg acc ccc gac gtg agc tac agc cag 1536
Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln
500 505 510

aag tac cgc gcc cgc atc cac tac gcc agc acc agc cag atc acc ttc 1584
Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe
515 520 525

acc ctg agc ctg gac ggg gcc ccc gct gca ccg ttc tac ttc gac aag 1632
Thr Leu Ser Leu Asp Gly Ala Pro Ala Ala Pro Phe Tyr Phe Asp Lys
530 535 540

acc atc aac aag ggc gac acc ctg acc tac aac agc ttc aac ctg gcc 1680
Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala
545 550 555 560

agc ttc agc acc cct ttc gag ctg agc ggc aac aac ctc cag atc ggc 1728
Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly
565 570 575

gtg acc ggc ctg agc gcc ggc gac aag gtg tac atc gac aag atc gag 1776
Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu
580 585 590

ttc atc ccc gtg aac tag 1794
Phe Ile Pro Val Asn
595

<210> 13
<211> 597
<212> PRT
<213> Secuencia Artificial

<220>
<221> misc_carácter
<222> (1609)..(1620)
<223> Secuencia codificadora del sitio de reconocimiento de Catepsina G

<400> 13

Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

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

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn
100 105 110

Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser
115 120 125

His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val
130 135 140

Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu
145 150 155 160

Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu
165 170 175

Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr
180 185 190

Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg
195 200 205

238
240

Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu
210 215 220

Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp
225 230 235 240

Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val
245 250 255

Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr
260 265 270

Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp
275 280 285

Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr
290 295 300

Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg
305 310 315 320

Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn
325 330 335

Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val
340 345 350

Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val
355 360 365

Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr
370 375 380

Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala
385 390 395 400

Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu
405 410 415

Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe
420 425 430

Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His
435 440 445

Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln
450 455 460

Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val
465 470 475 480

Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn
485 490 495

Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln
500 505 510

Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe
515 520 525

Thr Leu Ser Leu Asp Gly Ala Pro Ala Ala Pro Phe Tyr Phe Asp Lys
530 535 540

Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala
545 550 555 560

Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly
565 570 575

Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu
580 585 590

Phe Ile Pro Val Asn
595

<210> 14
<211> 1816
<212> ADN

<213> Secuencia Artificial

<220>
<221> CDS
<222> (1)..(1812)
<223> Secuencia codificadora cry3A058 modificada de maíz optimizado.

<220>
<221> misc_carácter
<222> (1621)..(1632)
<223> Secuencia codificadora de reconocimiento de Catepsina G
<400> 14
atg acg gcc gac aac aac acc gag gcc ctg gac agc agc acc acc aag 48
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

gac gtg atc cag aag ggc atc agc gtg gtg ggc gac ctg ctg ggc gtg 96
Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

gtg ggc ttc ccc ttc ggc ggc gcc ctg gtg agc ttc tac acc aac ttc 144
Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

ctg aac acc atc tgg ccc agc gag gac ccc tgg aag gcc ttc atg gag 192
Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

cag gtg gag gcc ctg atg gac cag aag atc gcc gac tac gcc aag aac 240
Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65 70 75 80

aag gca ctg gcc gag cta cag ggc ctc cag aac aac gtg gag gac tat 288
Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

gtg agc gcc ctg agc agc tgg cag aag aac ccc gtc tcg agc cgc aac 336
Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn
100 105 110

241
243

ccc cac agc cag ggc cgc atc cgc gag ctg ttc agc cag gcc gag agc 384
Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser
115 120 125

cac ttc cgc aac agc atg ccc agc ttc gcc atc agc ggc tac gag gtg 432
His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val
130 135 140

ctg ttc ctg acc acc tac gcc cag gcc gcc aac acc cac ctg ttc ctg 480
Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu
145 150 155 160

ctg aag gac gcc caa atc tac gga gag gag tgg ggc tac gag aag gag 528
Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu
165 170 175

gac atc gcc gag ttc tac aag cgc cag ctg aag ctg acc cag gag tac 576
Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr
180 185 190

acc gac cac tgc gtg aag tgg tac aac gtg ggt cta gac aag ctc cgc 624
Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg
195 200 205

ggc agc agc tac gag agc tgg gtg aac ttc aac cgc tac cgc cgc gag 672
Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu
210 215 220

atg acc ctg acc gtg ctg gac ctg atc gcc ctg ttc ccc ctg tac gac 720
Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp
225 230 235 240

gtg cgc ctg tac ccc aag gag gtg aag acc gag ctg acc cgc gac gtg 768
Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val
245 250 255

ctg acc gac ccc atc gtg ggc gtg aac aac ctg cgc ggc tac ggc acc 816
Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr
260 265 270

acc ttc agc aac atc gag aac tac atc cgc aag ccc cac ctg ttc gac 864
Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp
275 280 285

tac ctg cac cgc atc cag ttc cac acg cgt ttc cag ccc ggc tac tac 912
Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr
290 295 300

ggc aac gac agc ttc aac tac tgg agc ggc aac tac gtg agc acc cgc 960
Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg
305 310 315 320

ccc agc atc ggc agc aac gac atc atc acc agc ccc ttc tac ggc aac 1008
Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn
325 330 335

aag agc agc gag ccc gtg cag aac ctt gag ttc aac ggc gag aag gtg 1056
Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val
340 345 350

tac cgc gcc gtg gct aac acc aac ctg gcc gtg tgg ccc tct gca gtg 1104
Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val
355 360 365

tac agc ggc gtg acc aag gtg gag ttc agc cag tac aac gac cag acc 1152
Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr
370 375 380

gac gag gcc agc acc cag acc tac gac agc aag cgc aac gtg ggc gcc 1200
Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala
385 390 395 400

gtg agc tgg gac agc atc gac cag ctg ccc ccc gag acc acc gac gag 1248
Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu
405 410 415

ccc ctg gag aag ggc tac agc cac cag ctg aac tac gtg atg tgc ttc 1296
Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe
420 425 430

ctg atg cag ggc agc cgc ggc acc atc ccc gtg ctg acc tgg acc cac 1344
Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His
435 440 445

aag agc gtc gac ttc ttc aac atg atc gac agc aag aag atc acc cag 1392
Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln
450 455 460

ctg ccc ctg gtg aag gcc tac aag ctc cag agc ggc gcc agc gtg gtg 1440
Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val
465 470 475 480

gca ggc ccc cgc ttc acc ggc ggc gac atc atc cag tgc acc gag aac 1488
Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn
485 490 495

ggc agc gcc gcc acc atc tac gtg acc ccc gac gtg agc tac agc cag 1536
Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln
500 505 510

aag tac cgc gcc cgc atc cac tac gcc agc acc agc cag atc acc ttc 1584
Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe
515 520 525

acc ctg agc ctg gac ggg gcc ccc ttc aac caa tac gct gca ccg ttc 1632
Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Ala Ala Pro Phe
530 535 540

tac ttc gac aag acc atc aac aag ggc gac acc ctg acc tac aac agc 1680
Tyr Phe Asp Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser
545 550 555 560

ttc aac ctg gcc agc ttc agc acc cct ttc gag ctg agc ggc aac aac 1728
Phe Asn Leu Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn
565 570 575

ctc cag atc ggc gtg acc ggc ctg agc gcc ggc gac aag gtg tac atc 1776
Leu Gln Ile Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile
580 585 590

gac aag atc gag ttc atc ccc gtg aac tag atc tga gctc 1816
Asp Lys Ile Glu Phe Ile Pro Val Asn Ile
595 600

<210> 15
<211> 601
<212> PRT
<213> Secuencia Artificial

<220>
<221> misc_carácter
<222> (1621)..(1632)
<223> Secuencia codificadora del sitio de reconocimiento de Catepsina G
<400> 15
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

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

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn
100 105 110

Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser
115 120 125

His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val

~~245~~
247

130 135 140

Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu
145 150 155 160

Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu
165 170 175

Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr
180 185 190

Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg
195 200 205

Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu
210 215 220

Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp
225 230 235 240

Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val
245 250 255

Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr
260 265 270

Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp
275 280 285

Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr
290 295 300

Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg
305 310 315 320

Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn
325 330 335

Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val
340 345 350

Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val
355 360 365

Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr
370 375 380

Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala
385 390 395 400

Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu
405 410 415

Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe
420 425 430

Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His
435 440 445

Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln
450 455 460

Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val
465 470 475 480

Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn
485 490 495

Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln
500 505 510

Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe
515 520 525

Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Ala Ala Pro Phe
530 535 540

Tyr Phe Asp Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser
545 550 555 560

Phe Asn Leu Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn
565 570 575

Leu Gln Ile Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile
580 585 590

Asp Lys Ile Glu Phe Ile Pro Val Asn
595 600

<210> 16
<211> 1813
<212> ADN
<213> Secuencia Artificial

<220>
<221> CDS
<222> (1)..(1812)
<223> Secuencia codificadora de cry3A057 modificada optimizada en maíz.

<220>
<221> misc_carácter
<222> (322)..(333)
<223> Secuencia codificadora del sitio de reconocimiento de Catepsina G
<220>
<221> misc_carácter
<222> (1618)..(1629)
<223> Secuencia codificadora del sitio de reconocimiento de Catepsina G

<400> 16
atg acg gcc gac aac aac acc gag gcc ctg gac agc agc acc acc aag 48
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

gac gtg atc cag aag ggc atc agc gtg gtg ggc gac ctg ctg ggc gtg 96
Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

gtg ggc ttc ccc ttc ggc ggc gcc ctg gtg agc ttc tac acc aac ttc 144
Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

ctg aac acc atc tgg ccc agc gag gac ccc tgg aag gcc ttc atg gag 192
Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

cag gtg gag gcc ctg atg gac cag aag atc gcc gac tac gcc aag aac 240
Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65 70 75 80

aag gca ctg gcc gag cta cag ggc ctc cag aac aac gtg gag gac tat 288
Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

gtg agc gcc ctg agc agc tgg cag aag aac ccc gct gca ccg ttc ccc 336
Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Pro
100 105 110

cac agc cag ggc cgc atc cgc gag ctg ttc agc cag gcc gag agc cac 384
His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser His
115 120 125

ttc cgc aac agc atg ccc agc ttc gcc atc agc ggc tac gag gtg ctg 432
Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val Leu
130 135 140

ttc ctg acc acc tac gcc cag gcc gcc aac acc cac ctg ttc ctg ctg 480
Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu Leu
145 150 155 160

aag gac gcc caa atc tac gga gag gag tgg ggc tac gag aag gag gac 528
Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu Asp
165 170 175

atc gcc gag ttc tac aag cgc cag ctg aag ctg acc cag gag tac acc 576
Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr Thr
180 185 190

gac cac tgc gtg aag tgg tac aac gtg ggt cta gac aag ctc cgc ggc 624
Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg Gly
195 200 205

249

251

agc agc tac gag agc tgg gtg aac ttc aac cgc tac cgc cgc gag atg 672
Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu Met
210 215 220

acc ctg acc gtg ctg gac ctg atc gcc ctg ttc ccc ctg tac gac gtg 720
Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp Val
225 230 235 240

cgc ctg tac ccc aag gag gtg aag acc gag ctg acc cgc gac gtg ctg 768
Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val Leu
245 250 255

acc gac ccc atc gtg ggc gtg aac aac ctg cgc ggc tac ggc acc acc 816
Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr Thr
260 265 270

ttc agc aac atc gag aac tac atc cgc aag ccc cac ctg ttc gac tac 864
Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp Tyr
275 280 285

ctg cac cgc atc cag ttc cac acg cgt ttc cag ccc ggc tac tac ggc 912
Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr Gly
290 295 300

aac gac agc ttc aac tac tgg agc ggc aac tac gtg agc acc cgc ccc 960
Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg Pro
305 310 315 320

agc atc ggc agc aac gac atc atc acc agc ccc ttc tac ggc aac aag 1008
Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn Lys
325 330 335

agc agc gag ccc gtg cag aac ctt gag ttc aac ggc gag aag gtg tac 1056
Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val Tyr
340 345 350

cgc gcc gtg gct aac acc aac ctg gcc gtg tgg ccc tct gca gtg tac 1104
Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val Tyr
355 360 365

~~250~~
251

agc ggc gtg acc aag gtg gag ttc agc cag tac aac gac cag acc gac 1152
Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr Asp
370 375 380

gag gcc agc acc cag acc tac gac agc aag cgc aac gtg ggc gcc gtg 1200
Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala Val
385 390 395 400

agc tgg gac agc atc gac cag ctg ccc ccc gag acc acc gac gag ccc 1248
Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu Pro
405 410 415

ctg gag aag ggc tac agc cac cag ctg aac tac gtg atg tgc ttc ctg 1296
Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe Leu
420 425 430

atg cag ggc agc cgc ggc acc atc ccc gtg ctg acc tgg acc cac aag 1344
Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His Lys
435 440 445

agc gtc gac ttc ttc aac atg atc gac agc aag aag atc acc cag ctg 1392
Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln Leu
450 455 460

ccc ctg gtg aag gcc tac aag ctc cag agc ggc gcc agc gtg gtg gca 1440
Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val Ala
465 470 475 480

ggc ccc cgc ttc acc ggc ggc gac atc atc cag tgc acc gag aac ggc 1488
Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn Gly
485 490 495

agc gcc gcc acc atc tac gtg acc ccc gac gtg agc tac agc cag aag 1536
Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln Lys
500 505 510

tac cgc gcc cgc atc cac tac gcc agc acc agc cag atc acc ttc acc 1584
Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe Thr
515 520 525

ctg agc ctg gac ggg gcc ccc ttc aac caa tac gct gca ccg ttc tac 1632
Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Ala Ala Pro Phe Tyr
530 535 540

ttc gac aag acc atc aac aag ggc gac acc ctg acc tac aac agc ttc 1680
Phe Asp Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe
545 550 555 560

aac ctg gcc agc ttc agc acc cct ttc gag ctg agc ggc aac aac ctc 1728
Asn Leu Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu
565 570 575

cag atc ggc gtg acc ggc ctg agc gcc ggc gac aag gtg tac atc gac 1776
Gln Ile Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp
580 585 590

aag atc gag ttc atc ccc gtg aac tag atc tga gct c 1813
Lys Ile Glu Phe Ile Pro Val Asn Ile Ala
595 . 600

<210> 17
<211> 600
<212> PRT
<213> Secuencia Artificial

<220>
<221> misc_carácter
<222> (322)..(333)
<223> Secuencia codificadora del sitio de reconocimiento de Catepsina G

<220>
<221> misc_carácter
<222> (1618)..(1629)
<223> Secuencia codificadora del sitio de reconocimiento de Catepsina G

<400> 17
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

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

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Pro
100 105 110

His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser His
115 120 125

Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val Leu
130 135 140

Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu Leu
145 150 155 160

Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu Asp
165 170 175

Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr Thr
180 185 190

Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg Gly
195 200 205

Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu Met
210 215 220

285
255

Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp Val
225 230 235 240

Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val Leu
245 250 255

Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr Thr
260 265 270

Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp Tyr
275 280 285

Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr Gly
290 295 300

Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg Pro
305 310 315 320

Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn Lys
325 330 335

Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val Tyr
340 345 350

Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val Tyr
355 360 365

Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr Asp
370 375 380

Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala Val
385 390 395 400

Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu Pro
405 410 415

Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe Leu
420 425 430

Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His Lys

435 440 445

Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln Leu
450 455 460

Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val Ala
465 470 475 480

Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn Gly
485 490 495

Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln Lys
500 505 510

Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe Thr
515 520 525

Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Ala Ala Pro Phe Tyr
530 535 540

Phe Asp Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe
545 550 555 560

Asn Leu Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu
565 570 575

Gln Ile Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp
580 585 590

Lys Ile Glu Phe Ile Pro Val Asn
595 600

<210> 18
<211> 1819
<212> ADN
<213> Secuencia Artificial

<220>
<221> CDS

<222> (1)..(1818)
<223> Secuencia codificadora de cry3A056 optimizada de maíz.

<220>
<221> misc_carácter
<222> (322)..(333)
<223> Secuencia codificadora del sitio de reconocimiento de Catepsina G

<220>
<221> misc_carácter
<222> (1624)..(1635)
<223> Secuencia codificadora del sitio de reconocimiento de Catepsina G

<400> 18
atg acg gcc gac aac aac acc gag gcc ctg gac agc agc acc acc aag 48
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

gac gtg atc cag aag ggc atc agc gtg gtg ggc gac ctg ctg ggc gtg 96
Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

gtg ggc ttc ccc ttc ggc ggc gcc ctg gtg agc ttc tac acc aac ttc 144
Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

ctg aac acc atc tgg ccc agc gag gac ccc tgg aag gcc ttc atg gag 192
Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

cag gtg gag gcc ctg atg gac cag aag atc gcc gac tac gcc aag aac 240
Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65 70 75 80

aag gca ctg gcc gag cta cag ggc ctc cag aac aac gtg gag gac tat 288
Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

gtg agc gcc ctg agc agc tgg cag aag aac ccc gct gca ccg ttc cgc 336
Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Arg

100 105 110

aac ccc cac agc cag ggc cgc atc cgc gag ctg ttc agc cag gcc gag 384
Asn Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu
115 120 125

agc cac ttc cgc aac agc atg ccc agc ttc gcc atc agc ggc tac gag 432
Ser His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu
130 135 140

gtg ctg ttc ctg acc acc tac gcc cag gcc gcc aac acc cac ctg ttc 480
Val Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe
145 150 155 160

ctg ctg aag gac gcc caa atc tac gga gag gag tgg ggc tac gag aag 528
Leu Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys
165 170 175

gag gac atc gcc gag ttc tac aag cgc cag ctg aag ctg acc cag gag 576
Glu Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu
180 185 190

tac acc gac cac tgc gtg aag tgg tac aac gtg ggt cta gac aag ctc 624
Tyr Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu
195 200 205

cgc ggc agc agc tac gag agc tgg gtg aac ttc aac cgc tac cgc cgc 672
Arg Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg
210 215 220

gag atg acc ctg acc gtg ctg gac ctg atc gcc ctg ttc ccc ctg tac 720
Glu Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr
225 230 235 240

gac gtg cgc ctg tac ccc aag gag gtg aag acc gag ctg acc cgc gac 768
Asp Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp
245 250 255

gtg ctg acc gac ccc atc gtg ggc gtg aac aac ctg cgc ggc tac ggc 816
Val Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly

260 265 270

acc acc ttc agc aac atc gag aac tac atc cgc aag ccc cac ctg ttc 864
Thr Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe
275 280 285

gac tac ctg cac cgc atc cag ttc cac acg cgt ttc cag ccc ggc tac 912
Asp Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr
290 295 300

tac ggc aac gac agc ttc aac tac tgg agc ggc aac tac gtg agc acc 960
Tyr Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr
305 310 315 320

cgc ccc agc atc ggc agc aac gac atc atc acc agc ccc ttc tac ggc 1008
Arg Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly
325 330 335

aac aag agc agc gag ccc gtg cag aac ctt gag ttc aac ggc gag aag 1056
Asn Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys
340 345 350

gtg tac cgc gcc gtg gct aac acc aac ctg gcc gtg tgg ccc tct gca 1104
Val Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala
355 360 365

gtg tac agc ggc gtg acc aag gtg gag ttc agc cag tac aac gac cag 1152
Val Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln
370 375 380

acc gac gag gcc agc acc cag acc tac gac agc aag cgc aac gtg ggc 1200
Thr Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly
385 390 395 400

gcc gtg agc tgg gac agc atc gac cag ctg ccc ccc gag acc acc gac 1248
Ala Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp
405 410 415

gag ccc ctg gag aag ggc tac agc cac cag ctg aac tac gtg atg tgc 1296
Glu Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys

420 425 430

ttc ctg atg cag ggc agc cgc ggc acc atc ccc gtg ctg acc tgg acc 1344
Phe Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr
435 440 445

cac aag agc gtc gac ttc ttc aac atg atc gac agc aag aag atc acc 1392
His Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr
450 455 460

cag ctg ccc ctg gtg aag gcc tac aag ctc cag agc ggc gcc agc gtg 1440
Gln Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val
465 470 475 480

gtg gca ggc ccc cgc ttc acc ggc ggc gac atc atc cag tgc acc gag 1488
Val Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu
485 490 495

aac ggc agc gcc gcc acc atc tac gtg acc ccc gac gtg agc tac agc 1536
Asn Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser
500 505 510

cag aag tac cgc gcc cgc atc cac tac gcc agc acc agc cag atc acc 1584
Gln Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr
515 520 525

ttc acc ctg agc ctg gac ggg gcc ccc ttc aac caa tac gct gca ccg 1632
Phe Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Ala Ala Pro
530 535 540

ttc tac ttc gac aag acc atc aac aag ggc gac acc ctg acc tac aac 1680
Phe Tyr Phe Asp Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn
545 550 555 560

agc ttc aac ctg gcc agc ttc agc acc cct ttc gag ctg agc ggc aac 1728
Ser Phe Asn Leu Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn
565 570 575

aac ctc cag atc ggc gtg acc ggc ctg agc gcc ggc gac aag gtg tac 1776
Asn Leu Gln Ile Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr

580 585 590

atc gac aag atc gag ttc atc ccc gtg aac tag atc tga gct c 1819
Ile Asp Lys Ile Glu Phe Ile Pro Val Asn Ile Ala
595 600

<210> 19
<211> 602
<212> PRT
<213> Secuencia Artificial

<220>
<221> misc_carácter
<222> (322)..(333)
<223> Secuencia codificadora del sitio de reconocimiento de Catepsina G

<220>
<221> misc_carácter
<222> (1624)..(1635)
<223> Secuencia codificadora del sitio de reconocimiento de Catepsina G

<400> 19
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

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

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

~~260~~
262

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Arg
100 105 110

Asn Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu
115 120 125

Ser His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu
130 135 140

Val Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe
145 150 155 160

Leu Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys
165 170 175

Glu Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu
180 185 190

Tyr Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu
195 200 205

Arg Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg
210 215 220

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

Asp Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp
245 250 255

Val Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly
260 265 270

Thr Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe
275 280 285

Asp Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr
290 295 300

Tyr Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr
305 310 315 320

Arg Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly
325 330 335

Asn Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys
340 345 350

Val Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala
355 360 365

Val Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln
370 375 380

Thr Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly
385 390 395 400

Ala Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp
405 410 415

Glu Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys
420 425 430

Phe Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr
435 440 445

His Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr
450 455 460

Gln Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val
465 470 475 480

Val Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu
485 490 495

Asn Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser
500 505 510

Gln Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr

515 520 525

Phe Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Ala Ala Pro
530 535 540

Phe Tyr Phe Asp Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn
545 550 555 560

Ser Phe Asn Leu Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn
565 570 575

Asn Leu Gln Ile Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr
580 585 590

Ile Asp Lys Ile Glu Phe Ile Pro Val Asn
595 600

<210> 20
<211> 1797
<212> ADN
<213> Secuencia Artificial

<220>
<221> CDS
<222> (1)..(1791)
<223> Secuencia codificadora de cry3A063 modificado de maíz optimizado.

<220>
<221> misc_carácter
<222> (322)..(333)
<223> Secuencia codificadora del sitio de reconocimiento de Catepsina G.

<220>
<221> misc_carácter
<222> (1612)..(1623)
<223> Secuencia codificadora del sitio de reconocimiento de Catepsina G.

<400> 20
atg acg gcc gac aac aac acc gag gcc ctg gac agc agc acc acc aag 48

Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

gac gtg atc cag aag ggc atc agc gtg gtg ggc gac ctg ctg ggc gtg 96
Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

gtg ggc ttc ccc ttc ggc ggc gcc ctg gtg agc ttc tac acc aac ttc 144
Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

ctg aac acc atc tgg ccc agc gag gac ccc tgg aag gcc ttc atg gag 192
Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

cag gtg gag gcc ctg atg gac cag aag atc gcc gac tac gcc aag aac 240
Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65 70 75 80

aag gca ctg gcc gag cta cag ggc ctc cag aac aac gtg gag gac tat 288
Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

gtg agc gcc ctg agc agc tgg cag aag aac ccc gct gca ccg ttc cgc 336
Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Arg
100 105 110

aac ccc cac agc cag ggc cgc atc cgc gag ctg ttc agc cag gcc gag 384
Asn Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu
115 120 125

agc cac ttc cgc aac agc atg ccc agc ttc gcc atc agc ggc tac gag 432
Ser His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu
130 135 140

gtg ctg ttc ctg acc acc tac gcc cag gcc gcc aac acc cac ctg ttc 480
Val Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe
145 150 155 160

ctg ctg aag gac gcc caa atc tac gga gag gag tgg ggc tac gag aag 528

Leu Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys
165 170 175

gag gac atc gcc gag ttc tac aag cgc cag ctg aag ctg acc cag gag 576
Glu Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu
180 185 190

tac acc gac cac tgc gtg aag tgg tac aac gtg ggt cta gac aag ctc 624
Tyr Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu
195 200 205

cgc ggc agc agc tac gag agc tgg gtg aac ttc aac cgc tac cgc cgc 672
Arg Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg
210 215 220

gag atg acc ctg acc gtg ctg gac ctg atc gcc ctg ttc ccc ctg tac 720
Glu Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr
225 230 235 240

gac gtg cgc ctg tac ccc aag gag gtg aag acc gag ctg acc cgc gac 768
Asp Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp
245 250 255

gtg ctg acc gac ccc atc gtg ggc gtg aac aac ctg cgc ggc tac ggc 816
Val Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly
260 265 270

acc acc ttc agc aac atc gag aac tac atc cgc aag ccc cac ctg ttc 864
Thr Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe
275 280 285

gac tac ctg cac cgc atc cag ttc cac acg cgt ttc cag ccc ggc tac 912
Asp Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr
290 295 300

tac ggc aac gac agc ttc aac tac tgg agc ggc aac tac gtg agc acc 960
Tyr Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr
305 310 315 320

cgc ccc agc atc ggc agc aac gac atc atc acc agc ccc ttc tac ggc 1008

Arg Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly
325 330 335

aac aag agc agc gag ccc gtg cag aac ctt gag ttc aac ggc gag aag 1056
Asn Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys
340 345 350

gtg tac cgc gcc gtg gct aac acc aac ctg gcc gtg tgg ccc tct gca 1104
Val Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala
355 360 365

gtg tac agc ggc gtg acc aag gtg gag ttc agc cag tac aac gac cag 1152
Val Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln
370 375 380

acc gac gag gcc agc acc cag acc tac gac agc aag cgc aac gtg ggc 1200
Thr Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly
385 390 395 400

gcc gtg agc tgg gac agc atc gac cag ctg ccc ccc gag acc acc gac 1248
Ala Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp
405 410 415

gag ccc ctg gag aag ggc tac agc cac cag ctg aac tac gtg atg tgc 1296
Glu Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys
420 425 430

ttc ctg atg cag ggc agc cgc ggc acc atc ccc gtg ctg acc tgg acc 1344
Phe Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr
435 440 445

cac aag agc gtc gac ttc ttc aac atg atc gac agc aag aag atc acc 1392
His Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr
450 455 460

cag ctg ccc ctg gtg aag gcc tac aag ctc cag agc ggc gcc agc gtg 1440
Gln Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val
465 470 475 480

gtg gca ggc ccc cgc ttc acc ggc ggc gac atc atc cag tgc acc gag 1488

Val Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu
485 490 495

aac ggc agc gcc gcc acc atc tac gtg acc ccc gac gtg agc tac agc 1536
Asn Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser
500 505 510

cag aag tac cgc gcc cgc atc cac tac gcc agc acc agc cag atc acc 1584
Gln Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr
515 520 525

ttc acc ctg agc ctg gac ggg gcc ccc gct gca ccg ttc tac ttc gac 1632
Phe Thr Leu Ser Leu Asp Gly Ala Pro Ala Ala Pro Phe Tyr Phe Asp
530 535 540

aag acc atc aac aag ggc gac acc ctg acc tac aac agc ttc aac ctg 1680
Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu
545 550 555 560

gcc agc ttc agc acc cct ttc gag ctg agc ggc aac aac ctc cag atc 1728
Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile
565 570 575

ggc gtg acc ggc ctg agc gcc ggc gac aag gtg tac atc gac aag atc 1776
Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile
580 585 590

gag ttc atc ccc gtg aactag 1797
Glu Phe Ile Pro Val
595

<210> 21
<211> 597
<212> PRT
<213> Secuencia Artificial

<220>
<221> misc_carácter
<222> (322)..(333)

<223> Secuencia codificadora del sitio de reconocimiento de Catepsina G.
<220>
<221> misc_carácter
<222> (1612)..(1623)
<223> Secuencia codificadora del sitio de reconocimiento de Catepsina G.

<400> 21

Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

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

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Arg
100 105 110

Asn Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu
115 120 125

Ser His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu
130 135 140

Val Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe
145 150 155 160

Leu Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys
165 170 175

Glu Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu
180 185 190

Tyr Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu
195 200 205

Arg Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg
210 215 220

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

Asp Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp
245 250 255

Val Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly
260 265 270

Thr Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe
275 280 285

Asp Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr
290 295 300

Tyr Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr
305 310 315 320

Arg Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly
325 330 335

Asn Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys
340 345 350

Val Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala
355 360 365

Val Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln
370 375 380

Thr Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly
385 390 395 400

Ala Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp
405 410 415

Glu Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys
420 425 430

Phe Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr
435 440 445

His Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr
450 455 460

Gln Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val
465 470 475 480

Val Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu
485 490 495

Asn Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser
500 505 510

Gln Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr
515 520 525

Phe Thr Leu Ser Leu Asp Gly Ala Pro Ala Ala Pro Phe Tyr Phe Asp
530 535 540

Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu
545 550 555 560

Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile
565 570 575

Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile
580 585 590

Glu Phe Ile Pro Val

<210> 22
<211> 21
<212> ADN
<213> Secuencia Artificial

<220>
<221> misc_carácter
<222> (1)..(21)
<223> Cebador BamExt1
<400> 22
ggatccacca tgacggccga c 21

<210> 23
<211> 29
<212> ADN
<213> Secuencia Artificial

<220>
<221> misc_carácter
<222> (1)..(29)
<223> Cebador AAPFcola3
<400> 23
gaacggtgca gcggggttct tctgccagc 29

<210> 24
<211> 29
<212> ADN
<213> Secuencia Artificial

<220>
<221> misc_carácter
<222> (1)..(29)
<223> Cebador AAPFcola4

<400> 24

gctgcaccgt tcccccacag ccagggccg 29

<210> 25
<211> 21
<212> ADN
<213> Secuencia Artificial

<220>
<221> misc_carácter
<222> (1)..(21)
<223> Cebador XbalExt2

<400> 25
tctagaccca cgttgtacca c 21

<210> 26
<211> 29
<212> ADN
<213> Secuencia Artificial

<220>
<221> misc_carácter
<222> (1)..(29)
<223> Cebador Tail5mod

<400> 26
gctgcaccgt tccgcaaccc ccacagcca 29

<210> 27
<211> 19
<212> ADN
<213> Secuencia Artificial

<220>
<221> misc_carácter
<222> (1)..(19)
<223> Cebaor SalExt

<400> 27
gagcgtcgac ttcttcaac 19

<210> 28
<211> 30
<212> ADN
<213> Secuencia Artificial

<220>
<221> misc_carácter
<222> (1)..(30)
<223> Cebador AAPF-Y2

<400> 28
gaacggtgca gcgtattggt tgaaggggc 30

<210> 29
<211> 30
<212> ADN
<213> Secuencia Artificial

<220>
<221> misc_carácter
<222> (1)..(30)
<223> Cebador AAPF-Y1

<400> 29
gctgcaccgt tctacttcga caagaccatc 30

<210> 30
<211> 21
<212> ADN
<213> Secuencia Artificial

<220>
<221> misc_carácter

<222> (1)..(21)
<223> Cebador SacExt
<400> 30
gagctcagat ctagttcacg g 21

<210> 31
<211> 32
<212> ADN
<213> Secuencia Artificial

<220>
<221> misc_carácter
<222> (1)..(32)
<223> Cebador BBmod1
<400> 31
cggggccccc gctgcaccgt tctacttcga ca 32

<210> 32
<211> 32
<212> ADN
<213> Secuencia Artificial

<220>
<221> misc_carácter
<222> (1)..(32)
<223> Cebador BBmod2

<400> 32
tgtcgaagta gaacggtgca gcgggggccc cg 32

<210> 33
<211> 48
<212> ADN
<213> Secuencia Artificial

<220>
<221> misc_carácter

<222> (1)..(48)
<223> Cebador mo3Aext

<400> 33
ggatccacca tgaactacaa ggagttcctc cgcatgaccg ccgacaac 48

<210> 34
<211> 20
<212> ADN
<213> Secuencia Artificial

<220>
<221> misc_carácter
<222> (1)..(20)
<223> Cebador CMS16

<400> 34
cctccacctg ctccatgaag 20

PATENT COOPERATION TREATY

SEP 0 2 2003

From the:
INTERNATIONAL PRELIMINARY EXAMINING AUTHORITY

To:
HUSKISSON, Frank Mackie
SYNGENTA LIMITED
Intellectual Property Department
P.O. Box 3538
Bracknell
Berkshire, RG42 6YA
GRANDE BRETAGNE

VERIFY

PCT

WRITTEN OPINION

(PCT Rule 66)

Due: Nov. 20, 2003 Already Docketed 08

Date of mailing (day/month/year)	20.08.2003
Applicant's or agent's file reference	S-60065A
REPLY DUE	within 3 month(s) from the above date of mailing
International application No.	PCT/EP02/09789
International filing date (day/month/year)	02/09/2002
Priority date (day/month/year)	31/08/2001
International Patent Classification (IPC) or both national classification and IPC	C12N15/31
Applicant	SYNGENTA PARTICIPATIONS AG ✓

SYNGENTA CHEMICALS
ROA 00 NOV 03
22 AUG 2003
INVOICE ✓
PAGES

1. This written opinion is the first drawn up by this International Preliminary Examining Authority.

2. This opinion contains indications relating to the following items:

I	☒	Basis of the opinion
II	☐	Priority
III	☐	Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
IV	☐	Lack of unity of invention
V	☒	Reasoned statement under Rule 66.2(a)(ii) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
VI	☐	Certain document cited
VII	☐	Certain defects in the international application
VIII	☐	Certain observations on the international application

3. The applicant is hereby invited to reply to this opinion.

When? See the time limit indicated above. The applicant may, before the expiration of that time limit, request this Authority to grant an extension, see Rule 66.2(d).

How? By submitting a written reply, accompanied, where appropriate, by amendments, according to Rule 66.3. For the form and the language of the amendments, see Rules 66.8 and 66.9.

Also: For an additional opportunity to submit amendments, see Rule 66.4.
For the examiner's obligation to consider amendments and/or arguments, see Rule 66.4 bis.
For an informal communication with the examiner, see Rule 66.6.

If no reply is filed, the international preliminary examination report will be established on the basis of this opinion.

4. The final date by which the international preliminary examination report must be established according to Rule 69.2 is: 31/12/2003.

Name and mailing address of the international preliminary examining authority:
European Patent Office
D-80298 Munich
Tel. +49 89 2399 - 0 Tx: 523656 epmu d
Fax: +49 89 2399 - 4465

Authorized officer / Examiner
Paresce, D

Formalities officer (incl. extension of time limits)
Sülberg, A
Telephone No. +49 89 2399 7548

WRITTEN OPINION International application No. PCT/EP02/09789

I. Basis of the opinion

1. With regard to the **elements** of the international application (Replacement *sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this opinion as "originally filed"*):

Description, pages:

1-52 as originally filed

Claims, No.:

1-105 as originally filed

Sequence listing part of the description, pages:

1-62, as originally filed

2. With regard to the **language**, all the elements marked above were available or furnished to this Authority in the language in which the international application was filed, unless otherwise indicated under this item.

These elements were available or furnished to this Authority in the following language: , which is:

- ☐ the language of a translation furnished for the purposes of the international search (under Rule 23.1(b)).
- ☐ the language of publication of the international application (under Rule 48.3(b)).
- ☐ the language of a translation furnished for the purposes of international preliminary examination (under Rule 55.2 and/or 55.3).

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international preliminary examination was carried out on the basis of the sequence listing:

- ☒ contained in the international application in written form.
- ☒ filed together with the international application in computer readable form.
- ☐ furnished subsequently to this Authority in written form.
- ☐ furnished subsequently to this Authority in computer readable form.
- ☐ The statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.
- ☐ The statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished.

4. The amendments have resulted in the cancellation of:

- ☐ the description, pages:
- ☐ the claims, Nos.:

☐ the drawings, sheets:

5. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

   (Any replacement sheet containing such amendments must be referred to under item 1 and annexed to this report.)

6. Additional observations, if necessary:

V. Reasoned statement under Rule 66.2(a)(ii) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Claims	
Inventive step (IS)	Claims	1-105
Industrial applicability (IA)	Claims	

2. Citations and explanations
   see separate sheet

V. Reasoned statement under Rule 66.2(a)(ii) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1) The documents mentioned in this communication are numbered as in the search report, i.e. D1 corresponds to the first document of the International search report.

2) Inventive Step: Article 33(3) PCT

The subject-matter of claims 1-105 of the present application cannot be considered as involving an inventive step (Article 33(3) PCT) for the following reasons.

The documents D1 or D2 are regarded as being the closest prior art to the subject-matter of claims 1-105.

D1 discloses methods for increasing the activity of B. thuringiensis δ-endotoxins against Coleopteran insects (such as rootworm). Cry3A, Cry3B, and Cry3C toxins are described in D1. D1 discloses a method of improving the insecticidal activity of a B. thuringiensis Cry3Bb 6-endotoxin, which comprises inserting one or more protease sensitive sites into one or more loop regions of domain 1 of the 6-endotoxin (see p. 33). Two of the designed CryBb proteins (Cry3Bb.60 and Cry3Bb.11221) were engineered by identification and alteration of protease-sensitive sites and proteolytic processing. It is also mentioned in D1 that treatment of Cry3A toxin protein with trypsin, yields a stable cleavage product of 55 kDa from the 67 kDa native protein. The truncated Cry3A protein was found to retain the same level of insecticidal activity as the native protein. Cry3Bb toxin protein was also treated with trypsin. Unexpectedly, the bioactivity of the truncated Cry3Bb toxin was found to increase. Using this method, protease digestion of a B. thuringiensis toxin protein, a proteolytically sensitive site was identified on Cry3Bb, and a more highly active form of the protein (Cry3Bb. 60) was identified. Modifications to this proteolytically-sensitive site by introducing an additional protease recognition site also resulted in the isolation of a biologically more active protein (Cry3Bb. 11221). It is mentioned that Cry3Bb. 11221 is produced by B. thuringiensis as a full length toxin protein but is probably digested by insect gut

proteases to the same size as Cry3Bb. 60 (see example 7). The additional protease recognition site may make the la3,4 region even more sensitive to digestion, thereby increasing activity. D1 also mentions that proteolytically sensitive regions, once identified, may be modified or utilized to produce biologically more active toxins.

D2 also discloses methods for increasing the activity of B. thuringiensis δ-endotoxins against Coleopteran insects (such as rootworm). Isolated Cry3Bb polypeptides were modified to have improved insecticidal activity. Cry3B endotoxins, nucleic acid sequences and transgenic plants expressing said genes are disclosed. D2 discloses a method of improving the insecticidal activity of a Cry3Bb endotoxin, which comprises inserting one or more protease sensitive sites into one or more loop regions of domain 1 of the endotoxin. (see column 18, example 7).

D3 discloses modified Cry proteins, particularly modified CryIIIA proteins, having significantly altered toxicity, and DNA sequences encoding these proteins. Analysis of solvent-accessible amino acid positions in domain II of the CryIIIA protein by alanine-scanning mutagenesis identified individual amino acids involved in corn rootworm toxicity. The modified CryIIIA proteins AAW34811-41 have increased toxicity to Coleopteran insects (eg. corn rootworm).

D5 discloses that the insecticidal activity and specificity of different B. thuringiensis δ-endotoxins depends not only on the amino acid sequence of the endotoxin but also differential proteolytic processing within larval guts can determine the activity and specificity of the toxin (see p. 531 and 539)

The Examining Division is of the opinion that the addition of a protease recognition site to a CryA toxin is not inventive in view of D1 or D2. The documents disclose the addition of a protease recognition site to a CryB toxin. Introducing an additional protease recognition site resulted in the isolation of a biologically more active protein. It would be obvious to the person skilled in the art, namely when the same result is to be achieved, to apply the methods according to D1 or D2 with corresponding effect to a CryA toxin polypeptide as in D3 in which CryIIIA proteins are modified to have increased toxicity.

3) Clarity: Article 6 PCT

The application contains an unacceptable number of claims. Therefore, the requirements of Article 6 PCT regarding conciseness of claims has not been met and a detailed examination of the claims is difficult until a reasonable number of claims is on file. In this context, the attention of the applicant is drawn to the fact that undue repetition of wording between one claim and another should be avoided by use of the dependent form and that the relationship between the claims should be clearly visible.

Should the applicant file a new set of claims, which take account of the above comments, he is requested to clearly identify the amendments carried out, no matter whether they concern amendments by addition, replacement or deletion, and to indicate the passages of the application as filed on which these amendments are based (see Rule 66.8(a) PCT), to facilitate the examination of the conformity of the amended application with the requirements of Article 34(2)(b) PCT. If the applicant regards it as appropriate these indications could be submitted in handwritten form on a copy of the relevant parts of the application as filed. The attention of the Applicant is drawn to the fact that the application may not be amended in such a way that it contains subject-matter which extends beyond the content of the application as filed, Article 34(2)(b) PCT.

TRATADO DE COOPERACION EN MATERIA DE PATENTES

PCT

INFORME DE BÚSQUEDA INTERNACIONAL

(Artículo 18 y Reglas 43 Y 44 del PCT)

Referencia del expediente del solicitante o del mandatario **S-60065A**	**PARA ACCIÓN** véase la notificación de transmisión del informe de búsqueda internacional (formulario PCT/ISA/220) y, si procede, el punto 5 más abajo.	
Solicitud Internacional No. **PCT/EP 02/ 09789**	Fecha de presentación internacional *(día/mes/año)* **02/09/2002**	Fecha de prioridad *(más antigua) (día/mes/año)* **31/08/2001**
Solicitante **SYNGENTA PARTICIPATIONS AG**		

El presente informe de búsqueda internacional, emitido por la Administración encargada de la búsqueda internacional, se transmite al solicitante conforme al Artículo 18. Se envía una copia a la Oficina Internacional.

Este informe de búsqueda internacional comprende __4___ hojas.

[X] Está acompañado de una copia de cada documento relativo al estado de la técnica citado en el mismo.

1. **Base del Informe**
   a. Por lo que respecta al idioma, se ha efectuado la búsqueda internacional sobre la base de la solicitud internacional en el idioma en el que fue presentada, salvo que se indique lo contrario en el mismo punto

   [ ] se ha efectuado la búsqueda internacional sobre la base de una traducción de la solicitud internacional entregada a la Administración.

   b. Por lo que respecta a las secuencias de nucleótidos y/o aminoácidos divulgadas en la solicitud internacional (si procede), se ha efectuado la búsqueda internacional sobre la base de la lista de secuencias:

   [X] contenida en la solicitud internacional, en forma escrita.

   [X] presentada con la solicitud internacional, en forma legible por ordenador.

   [ ] entregada posteriormente a la Administración, en forma escrita.

   [ ] entregada posteriormente a la Administración, en forma legible por ordenador

   [ ] Ha sido entregada la declaración, según la cual la lista de secuencias presentada por escrito y entregada posteriormente no va más allá de la divulgación contenida en la solicitud tal como fue presentada.

   [ ] Ha sido entregada la declaración, según la cual las informaciones grabadas en forma legible por ordenador son idénticas a las de la lista de secuencias presentada por escrito.

2. [ ] **Se ha estimado que ciertas reivindicaciones no podían ser objeto de búsqueda** (véase el recuadro I).

3. [ ] **Existe falta de unidad de la invención** (véase el recuadro II).

4. Por lo que respecta al título,

   [X] se ha aprobado el texto tal como ha sido entregado por el solicitante.

   [ ] El texto ha sido redactado por la Administración y tiene el contenido siguiente:

5. Por lo que respecta al resumen,

   [X] se ha aprobado el texto tal como ha sido entregado por el solicitante.

   [ ] el texto (reproducido en el recuadro III) ha sido redactado por la Administración conforme a la Regla 38.2.b). El solicitante puede presentar observaciones a la Administración en el plazo de un mes desde la fecha de expedición del presente informe de búsqueda internacional.

6. La figura de los dibujos a publicar con el resumen es la Nº _____

   [ ] sugerida por el solicitante.

   [ ] porque el solicitante no ha sugerido figura.

   [ ] porque esa figura caracteriza mejor la invención.

   [ ] No hay ninguna figura que publicar.

INFORME DE BÚSQUEDA INTERNACIONAL

Solicitud internacional Nº
PCT/EP 02/09789

A. CLASIFICACIÓN DEL OBJETO DE LA SOLICITUD

IPC 7 C12N15/31 C12N15/82 C07K14/325 A01H5/00

Según la Clasificación Internacional de Patentes (IPC) o a la vez según la clasificación nacional y la IPC

B. CAMPOS SOBRE LOS QUE SE HA EFECTUADO LA BÚSQUEDA

Documentación mínima consultada (sistema de clasificación seguido de los símbolos de clasificación)

IPC 7 C07K

Documentación consultada distinta de la documentación mínima en la medida en que esos documentos correspondan a campos sobre los que se ha efectuado la búsqueda

Base de datos electrónica consultada durante la búsqueda internacional (nombre de la base de datos y, si fuera posible, términos de búsqueda utilizados)

EPO-Internal, WPI Data, BIOSIS, EMBASE, MEDLINE

C. DOCUMENTOS CONSIDERADOS PERTINENTES

Categoría*	Documentos citados con indicación, si procede, de los pasajes pertinentes	Nº de las reivindicaciones
Y	WO 99 31248 A (WALTERS FREDERICK S ; ENGLISH LEIGH (US); MONSANTO CO (US); BRYSON) 24 de junio 1999 (1999-06-24) ver reivindicaciones 115-6, ejemplo 7, y p. 33 —	1-105
Y	US 6 063 597 A (WALTERS FREDERICK S ET AL.) 16 de mayo 2000 (2000-05-16) ver columna 18, ejemplo 7 —	1-105
Y	US 5 659 123 A (VAN RIE JEROEN ET AL.) 19 de agosto 1997 (1997-08-19) ver columnas 2, 5 —	1-105
Y	WO 99 00407 A (RIE JEROEN VAN; ROECK SANDRA DE (BE); PLANT GENETIC SYSTEMS NV (BE) 7 enero 1999 (1999-01-07) ver p. 7 —	1-105

-/-

[X] Véase la continuación del recuadro C para el final de la lista de documentos

[X] Los documentos de familias de patentes se indican en anexo.

* Categorías especiales de documentos citados:

"A" documento que define el estado general de la técnica, no considerado como particularmente pertinente

"E" solicitud o patente anterior, pero publicada en la fecha de presentación internacional o después de esa fecha

"L" documento que puede plantear dudas sobre una reivindicación de prioridad o citado para determinar la fecha de publicación de otra cita o por una razón especial (tal como se indica)

"O" documento que se refiere a una divulgación oral, un uso, una exposición o cualquier otro medio

"P" documento publicado antes de la fecha de presentación internacional, pero después de la fecha de prioridad reivindicada

"T" documento ulterior publicado después de la fecha de presentación internacional o de prioridad y que no pertenece al estado de la técnica pertinente, pero citado para poder comprender el principio o la teoría que constituye la base de la invención

"X" documento particularmente pertinente; la invención reivindicada no puede considerarse nueva o que implica actividad inventiva en relación con el documento considerado aisladamente

"Y" documento particularmente pertinente; la invención reivindicada no puede considerarse que implica actividad inventiva cuando el documento está asociado a uno o varios otros documentos de la misma naturaleza, siendo esta combinación evidente para un experto en la materia

"&" documento que forma parte de la misma familia de patentes

Fecha en la que se ha finalizado efectivamente la búsqueda	Fecha de expedición del informe de búsqueda
19 de mayo 2003	28/05/2003
Nombre y dirección postal de la Administración encargada de la búsqueda internacional Oficina de Patentes Europea, P.B. 5818 Patentlaan 2 NL-2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Funcionario autorizado Paresce, D

INFORME DE BÚSQUEDA INTERNACIONAL

Solicitud internacional Nº
PCT/EP 02/09789

C (continuación). DOCUMENTOS CONSIDERADOS PERTINENTES

Categoría*	Documentos citados con indicación de los pasajes pertinentes, si procede	Nº de las reivindicaciones
Y	HAIDER M Z ET AL: "SPECIFIC OF BACILLUS THURINGIENSIS VAR. COLMERI INSECTICIDAL DELTA-ENDOTOXIN IN DETERMINED BY DIFFERENTIAL PROTEOLYTIC PROCESSING OF THE PROTOXIN BY LARVAL GUT PROTEASES" EUROPEAN JOURNAL OF BIOCHEMISTRY, BERLIN, DE vol. 156, no. 3, 1 mayo 1986 (1986-05-01), páginas 531-540, XP000567680 ISSN: 0014-2956 ver resumen, p. 1620-2 —	1-105
A	SMEDLEY ET AL: "Mutagenesis of three surface-exposed loops of a Bacillus thuringiensis insecticidal toxin reveals residues important for toxicidity, receptor recognition and possibly membrane insertion" MICROBIOLOGY, SOCIETY FOR GENERAL MICROBIOLOGY, READING, GB, vol. 142, no. PART 7, julio 1996 (1996-07), páginas 1617-1624, XP002103661 ISSN: 1350-0872 todo el documento —	1-105

INFORME DE BÚSQUEDA INTERNACIONAL

Solicitud internacional Nº
PCT/EP 02/09789

Documento de patente citado en el informe de búsqueda		Fecha de publicación	Miembro(s) de familias de patentes			Fecha de publicación
WO 9931248	A	24-06-1999	US	6063597	A	16-05-2000
			US	6060594	A	09-05-2000
			US	6077824	A	20-06-2000
			US	6023013	A	08-02-2000
			AU	2001399	A	05-07-1999
			BR	9814294	A	23-10-2001
			CA	2314429	A1	24-06-1999
			CN	1286728	T	07-03-2001
			EP	1040192	A1	04-10-2000
			WO	9931248	A1	24-06-1999
			ZA	9811673	A	28-06-1999
US 6063597	A	16-05-2000	AU	2001399	A	05-07-1999
			BR	9814294	A	23-10-2001
			CA	2314429	A1	24-06-1999
			CN	1286728	T	07-03-2001
			EP	1040192	A1	04-10-2000
			WO	9931248	A1	24-06-1999
			ZA	9811673	A	28-06-1999
US 5659123	A	19-08-1997	NINGUNO			
WO 9900407	A	07-01-1999	AU	741600	B2	06-12-2001
			AU	8804498	A	19-01-1999
			WO	9900407	A2	07-01-1999
			EP	0989998	A1	05-04-2000

S_60065 P1 285 287

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

June 26, 2002

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE UNDER 35 USC 111.

APPLICATION NUMBER: ***60/316,421***
FILING DATE: ***August 31, 2001***

By Authority of the
COMMISSIONER OF PATENTS AND TRADEMARKS

N. WOODSON
Certifying Officer

090401

APPROV Box 288

Docket Number	S-60086P1

FILING BY "EXPRESS MAIL" UNDER 37 CFR 1.10

ET482519002US	August 31, 2001
Express Mail Label Number	Date of Deposit

Address to: Commissioner for Patents
Box Provisional Patent Application
Washington, DC 20231

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

Transmitted herewith for filing under 37 CFR §1.53(c) is the PROVISIONAL APPLICATION for patent of

INVENTOR(S)		
Given Name (first and middle (if any))	Family Name or Surname	Residence (City and either State or Foreign Country)
Eric Cheryl	Chen Stacy	Chapel Hill, North Carolina Cary, North Carolina

TITLE OF THE INVENTION (280 characters max)

MODIFIED CRY3A TOXINS AND NUCLEIC ACID SEQUENCES CODING THEREFOR

CORRESPONDENCE ADDRESS

Direct all correspondence to the address associated with Customer No. 022647, which is currently:
022647
Syngenta Biotechnology Inc.
Patent Department
P.O. Box 12257
Research Triangle Park, NC 27709-2257

ENCLOSED APPLICATION PARTS *(check all that apply)*

- ☒ Specification (including Any Claims and Abstract) - 82 pages
- ☐ Drawings - sheets
- ☒ Other (specify): Sequence Listing, Sequence Diskette, verified statement

METHOD OF PAYMENT

The Commissioner is hereby authorized to charge any additional fees required to Deposit Account Number: 50-1744 in the name of Syngenta.	PROVISIONAL FILING FEE AMOUNT: $ 150

☐ U.S. Government agency and contract number: (If the invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.)

Respectfully submitted,

Gregory W. Warren

Date: August 31, 2001

Gregory W. Warren
Agent for Applicants
Reg. No. P-48,366
Tel. No. (919) 541-8646

60316421 083101

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

IN RE APPLICATION OF
CHEN ET AL.
APPLICATION NO: TBA
FILED: On Even Date
FOR: MODIFIED CRY3A TOXINS AND NUCLEIC ACID SEQUENCES CODING THEREFOR

ART UNIT: TBA
EXAMINER: TBA

Commissioner for Patents
Washington, D.C. 20231

SUBMISSION OF SEQUENCE LISTING
INCLUDING STATEMENT OF VERIFICATION

Sir:

Applicants hereby provide a Computer Readable Form of the Sequence Listing as well as the Paper Copy thereof. The undersigned states that the Paper Copy and the Computer Readable Form, submitted in accordance with 37 CFR §1.821(c) and (e), respectively, are the same.

Respectfully submitted,

Gregory W. Warren
Gregory W. Warren
Agent for Applicants
Reg. No. P-48,385

Syngenta Biotechnology, Inc.
Patent Department
P.O. Box 12257
Research Triangle Park, NC 27709-2257
(919) 541-8587
Date: August 31, 2001

021842.08101

185
290

Modified Cry3A Toxins and Nucleic Acid Sequences Coding Therefor

FIELD OF THE INVENTION

[0001] The present invention relates to the fields of protein engineering, plant molecular biology and pest control. More particularly, the present invention relates to novel modified Cry3A toxins and nucleic acid sequences whose expression results in the modified Cry3A toxins, and methods of making and methods of using the modified Cry3A toxins and corresponding nucleic acid sequences to control insects.

BACKGROUND OF THE INVENTION

[0002] Species of corn rootworm are considered to be the most destructive corn pests. In the United States the three important species are *Diabrotica virgifera virgifera*, the western corn rootworm; *D. longicornis barberi*, the northern corn rootworm and *D. undecimpunctata howardi*, the southern corn rootworm. Only western and northern corn rootworms are considered primary pests of corn in the US Corn Belt. Corn rootworm larvae cause the most substantial plant damage by feeding almost exclusively on corn roots. This injury has been shown to increase plant lodging, to reduce grain yield and vegetative yield as well as alter the nutrient content of the grain. Larval feeding also causes indirect effects on maize by opening avenues through the roots for bacterial and fungal infections which lead to root and stalk rot diseases. Adult corn rootworms are active in cornfields in late summer where they feed on ears, silks and pollen, interfering with normal pollination.

[0003] Corn rootworms are mainly controlled by intensive applications of chemical pesticides, which are active through inhibition of insect growth, prevention of insect feeding or reproduction, or cause death. Good corn rootworm control can thus be reached, but these chemicals can sometimes also affect other, beneficial organisms. Another problem resulting from the wide use of chemical pesticides is the appearance of resistant insect varieties. Yet another problem is due to the fact that corn rootworm larvae feed underground thus making it difficult to apply rescue treatments of insecticides. Therefore, most insecticide applications are made prophylactically at the time of planting. This practice results in a large environmental burden. This has been partially alleviated by

EP 1 037 981 B1 2003-10-01

285
291

various farm management practices, but there is an increasing need for alternative pest control mechanisms.

[0004] Biological pest control agents, such as *Bacillus thuringiensis* (Bt) strains expressing pesticidal toxins like δ-endotoxins, have also been applied to crop plants with satisfactory results against primarily lepidopteran insect pests. The δ-endotoxins are proteins held within a crystalline matrix that are known to possess insecticidal activity when ingested by certain insects. The various δ-endotoxins have been classified based upon their spectrum of activity and sequence homology. Prior to 1990, the major classes were defined by their spectrum of activity with the Cry1 proteins active against Lepidoptera (moths and butterflies), Cry2 proteins active against both Lepidoptera and Diptera (flies and mosquitoes), Cry3 proteins active against Coleoptera (beetles) and Cry4 proteins active against Diptera (Hofte and Whitely, 1989, Microbiol. Rev. 53:242-255). Recently a new nomenclature was developed which systematically classifies the Cry proteins based on amino acid sequence homology rather than insect target specificities (Crickmore *et al.* 1998, Microbiol. Molec. Biol. Rev. 62:807-813).

[0005] The spectrum of insecticidal activity of an individual δ-endotoxin from Bt is quite narrow, with a given δ-endotoxin being active against only a few species within an Order. For instance, the Cry3A protein is known to be very toxic to the Colorado potato beetle, *Leptinotarsa decemlineata*, but has very little or no toxicity to related beetles in the genus *Diabrotica* (Johnson *et al.*, 1993, J. Econ. Entomol. 86:330-333). According to Slaney *et al.* (1992, Insect Biochem. Molec. Biol. 22:9-18) the Cry3A protein is at least 2000 times less toxic to southern corn rootworm larvae than to the Colorado potato beetle. It is also known that Cry3A has little or no toxicity to the western corn rootworm.

[0006] Specificity of the δ-endotoxins is the result of the efficiency of the various steps involved in producing an active toxin protein and its subsequent interaction with the epithelial cells in the insect mid-gut. To be insecticidal, most known δ-endotoxins must first be ingested by the insect and proteolytically activated to form an active toxin. Activation of the insecticidal crystal proteins is a multi-step process. After ingestion, the crystals must first be solubilized in the insect gut. Once solubilized, the δ-endotoxins are activated by specific proteolytic cleavages. The proteases in the insect gut can play a role in specificity by determining where the δ-endotoxin is processed. Once the δ-endotoxin has been solubilized and processed it binds to specific receptors on the surface of the

390
292

insects' mid-gut epithelium and subsequently integrates into the lipid bilayer of the brush border membrane. Ion channels then form disrupting the normal function of the midgut eventually leading to the death of the insect.

[0007] In Lepidoptera, gut proteases process δ-endotoxins from 130-140 kDa protoxins to toxic proteins of approximately 60-70 kDa. Processing of the protoxin to toxin has been reported to proceed by removal of both N- and C-terminal amino acids with the exact location of processing being dependent on the specific insect gut fluids involved (Ogiwara *et al.*, 1992, J. Invert. Pathol. 60:121-126). The proteolytic activation of a δ-endotoxin can play a significant role in determining its specificity. For example, a δ-endotoxin from Bt var. *aizawa*, called IC1, has been classified as a Cry1Ab protein based on its sequence homology with other known Cry1Ab proteins. Cry1Ab proteins are typically active against lepidopteran insects. However, the IC1 protein has activity against both lepidopteran and dipteran insects depending upon how the protein is processed (Haider *et al.* 1986, Euro. J. Biochem. 156: 531-540). In a dipteran gut, a 53 kDa active IC1 toxin is obtained, whereas in a lepidopteran gut, a 55 kDa active IC1 toxin is obtained. IC1 differs from the holotype HD-1 Cry1Ab protein by only four amino acids, so gross changes in the receptor binding region do not seem to account for the differences in activity. The different proteolytic cleavages in the two different insect guts possibly allow the activated molecules to fold differently thus exposing different regions capable of binding different receptors. The specificity therefore, appears to reside with the gut proteases of the different insects.

[0008] Coleopteran insects have guts that are more neutral to acidic and coleopteran-specific δ-endotoxins are similar to the size of the activated lepidopteran-specific toxins. Therefore, the processing of coleopteran-specific δ-endotoxins was formerly considered unnecessary for toxicity. However, recent data suggests that coleopteran-active δ-endotoxins are solubilized and proteolyzed to smaller toxic polypeptides. The 73 kDa Cry3A δ-endotoxin protein produced by *B. thuringiensis* var. *tenebrionis* is readily processed in the bacterium at the N-terminus, losing 49-57 residues during or after crystal formation to produce the commonly isolated 67 kDa form (Carroll *et al.*, 1989, Biochem. J. 261:99-105). McPherson *et al.*, 1988 (Biotechnology 6:61-66) also demonstrated that the native *cry*3A gene contains two functional translational initiation codons in the same reading frame, one coding for the 73 kDa protein and the other coding for the 67 kDa

SO2154210.031101

protein starting at Met-1 and Met-48 respectively, of the deduced amino acid sequence (See SEQ ID NO: 2). Both proteins then can be considered naturally occurring full-length Cry3A proteins. Treatment of soluble 67 kDa Cry3A protein with either trypsin or insect gut extract results in a cleavage product of 55 kDa with Asn-159 of the deduced amino acid sequence at the N-terminus. This polypeptide was found to be as toxic to a susceptible coleopteran insect as the native 67 kDa Cry3A toxin. (Carroll *et al. Ibid*). Thus, a natural trypsin recognition site exists between Arg-158 and Asn-159 of the deduced amino acid sequence of the native Cry3A toxin (SEQ ID NO: 2). Cry3A can also be cleaved by chymotrypsin, resulting in three polypeptides of 49, 11, and 6 kDa. N-terminal analysis of the 49 and 6 kDa components showed the first amino acid residue to be Ser-162 and Tyr-588, respectively (Carroll *et al.*, 1997 J. Invert. Biol. 70:41-49). Thus, natural chymotrypsin recognition sites exist in Cry3A between His-161 and Ser-162 and between Tyr-587 and Tyr-588 of the deduced amino acid sequence (SEQ ID NO: 2). The 49 kDa chymotrypsin product appears to be more soluble at neutral pH than the native 67 kDa protein or the 55 kDa trypsin product and retains full insecticidal activity against the Cry3A-susceptible insects, Colorado potato beetle and mustard beetle, (*Phaedon cochleariae*).

[0009] Insect gut proteases typically function in aiding the insect in obtaining needed amino acids from dietary protein. The best understood insect digestive proteases are serine proteases that appear to be the most common (Englemann and Geraerts, 1980, J. Insect Physiol. 261:703-710), particularly in lepidopteran species. The majority of coleopteran larvae and adults, for example Colorado potato beetle, have slightly acidic midguts, and cysteine proteases provide the major proteolytic activity (Wolfson and Mudock, 1990, J. Chem. Ecol. 16:1089-1102). More precisely, Thie and Houseman (1990, Insect Biochem. 20:313-318) identified and characterized the cysteine proteases, cathepsin B and H, and the aspartyl protease, cathepsin D in Colorado potato beetle. Gillikin *et al.* (1992, Arch. Insect Biochem. Physiol. 19:285-298) characterized the proteolytic activity in the guts of western corn rootworm larvae and found 15, primarily cysteine, proteases. Until disclosed in this invention, no reports have indicated that the serine protease, cathepsin G, exists in western corn rootworm. The diversity and different activity levels of the insect gut proteases may influence an insect's sensitivity to a particular Bt toxin.

292
294

[0010] Many new and novel Bt strains and δ-endotoxins with improved or novel biological activities have been described over the past five years including strains active against nematodes (EP 0517367A1). However, relatively few of these strains and toxins have activity against coleopteran insects. Further, none of the now known coleopteran-active δ-endotoxins, for example Cry3A, Cry3B, Cry3C, Cry7A, Cry8A, Cry8B, and Cry8C, have sufficient oral toxicity against corn rootworm to provide adequate field control if delivered, for example, through microbes or transgenic plants. Therefore, other approaches for producing novel toxins active against corn rootworm need to be explored.

[0011] As more knowledge has been gained as to how the δ-endotoxins function, attempts to engineer δ-endotoxins to have new activities have increased. Engineering δ-endotoxins was made more possible by the solving of the three dimensional structure of Cry3A in 1991 (Li *et al.*, 1991, Nature 353:815-821). The protein has three structural domains: the N-terminal domain I, from residues 1-290, consists of 7 alpha helices, domain II, from residues 291-500, contains three beta-sheets and the C-terminal domain III, from residues 501-644, is a beta-sandwich. Based on this structure, a hypothesis has been formulated regarding the structure/function relationship of the δ-endotoxins. It is generally thought that domain I is primarily responsible for pore formation in the insect gut membrane (Gazit and Shai, 1993, Appl. Environ. Microbiol. 57:2816-2820), domain II is primarily responsible for interaction with the gut receptor (Ge *et al.*, 1991, J. Biol. Chem. 32:3429-3436) and that domain III is most likely involved with protein stability (Li *et al.* 1991, supra) as well as having a regulatory impact on ion channel activity (Chen *et al.*, 1993, PNAS 90:9041-9045).

[0012] Lepidopteran-active δ-endotoxins have been engineered in attempts to improve specific activity or to broaden the spectrum of insecticidal activity. For example, the silk moth (*Bombyx mori*) specificity domain from Cry1Aa was moved to Cry1Ac, thus imparting a new insecticidal activity to the resulting chimeric protein (Ge *et al.* 1989, PNAS 86: 4037-4041). Also, Bosch *et al.* 1998 (US Patent 5,736,131), created a new lepidopteran-active toxin by substituting domain III of Cry1B with domain III of Cry1C thus producing a Cry1B-Cry1C hybrid toxin with a broader spectrum of lepidopteran activity.

[0013] Several attempts at engineering the coleopteran-active δ-endotoxins have been reported. Van Rie *et al.*, 1997, (US Patent No. 5,659,123) engineered Cry3A by randomly

295
295

03364421 - 093101

replacing amino acids, thought to be important in solvent accessibility, in domain II with the amino acid alanine. Several of these random replacements confined to receptor binding domain II were reportedly involved in increased western corn rootworm toxicity. However, others have shown that some alanine replacements in domain II of Cry3A result in disruption of receptor binding or structural instability (Wu and Dean, 1996, J. Mol. Biol. 255: 628-640). English *et al.*, 1999, (Intl. Pat. Appl. Publ. No. WO 99/31248) reported amino acid substitutions in Cry3Bb that caused increases in toxicity to southern and western corn rootworm. However, of the 35 reported Cry3Bb mutants, only three, with mutations primarily in domain II and the domain II-domain I interface, were active against western corn rootworm. Further, the differences in toxicity of wild-type Cry3Bb against western corn rootworm in the same assays were greater than any of the differences between the mutated Cry3Bb toxins and the wild-type Cry3Bb. Therefore, improvements in toxicity of the Cry3Bb mutants appear to be confined primarily to southern corn rootworm.

[0014] There remains a need to design new and effective pest control agents that provide an economic benefit to farmers and that are environmentally acceptable. Particularly needed are modified Cry3A toxins that control western corn rootworm, the major pest of corn in the United States, that are or could become resistant to existing insect control agents. Furthermore, agents whose application minimizes the burden on the environment, as through transgenic plants, are desirable.

SUMMARY

[0015] In view of these needs, it is an object of the present invention to provide novel nucleic acid sequences encoding modified Cry3A toxins having increased toxicity to corn rootworm. By inserting a protease recognition site that is recognized by a target-insect gut protease in at least one position of a Cry3A toxin, in accordance with the present invention, a modified Cry3A toxin having significantly greater toxicity, particularly to western and northern corn rootworm is designed. The invention is further drawn to the novel modified Cry3A toxins resulting from the expression of the nucleic acid sequences, and to compositions and formulations containing the modified Cry3A toxins, which are capable of inhibiting the ability of insect pests to survive, grow and reproduce, or of limiting insect-related damage or loss to crop plants. The invention is further drawn to a

394
396

method of making the modified Cry3A toxins and to methods of using the modified *cry*3A nucleic acid sequences, for example in microorganisms to control insects or in transgenic plants to confer protection from insect damage, and to a method of using the modified Cry3A toxins, and compositions and formulations comprising the modified Cry3A toxins, for example applying the modified Cry3A toxins or compositions or formulations to insect-infested areas, or to prophylactically treat insect-susceptible areas or plants to confer protection against the insect pests.

[0016] The novel modified Cry3A toxins described herein are highly active against insects. For example, the modified Cry3A toxins of the present invention can be used to control economically important insect pests such as western corn rootworm (*Diabrotica virgifera virgifera*) and northern corn rootworm (*D. longicornis barberi*). The modified Cry3A toxins can be used singly or in combination with other insect control strategies to confer maximal pest control efficiency with minimal environmental impact.

[0017] According to one aspect, the present invention provides an isolated nucleic acid molecule comprising a nucleotide sequence that encodes a modified Cry3A toxin, wherein the modified Cry3A toxin comprises at least one additional protease recognition site that does not naturally occur in a Cry3A toxin. The additional protease recognition site, which is recognized by a gut protease of a target insect, is inserted at approximately the same position as a naturally occurring protease recognition site in the Cry3A toxin. The modified Cry3A toxin causes higher mortality to a target insect than the mortality caused by a Cry3A toxin to the same target insect. Preferably, the modified Cry3A toxin causes at least about 50 % mortality to a target insect to which a Cry3A toxin causes only up to about 30% mortality.

[0018] In one embodiment of this aspect, the gut protease of a target insect is selected from the group consisting of serine proteases, cysteine proteases and aspartic proteases. Preferable serine proteases according to this embodiment include cathepsin G, trypsin, chymotrypsin, carboxypeptidase, endopeptidase and elastase, most preferably cathepsin G.

[0019] In another embodiment of this aspect, the additional protease recognition site is inserted in either domain I or domain III or in both domain I and domain III of the Cry3A toxin. Preferably, the additional protease recognition site is inserted in either domain I or domain III or in both domain I and domain III at a position that replaces, is adjacent to, or is within a naturally occurring protease recognition site.

295
297

[0020] In a yet another embodiment, the additional protease recognition site is inserted in domain I between amino acids corresponding to amino acid numbers 154 and 162 of SEQ ID NO: 2. Preferably, the additional protease recognition site is inserted between amino acid numbers 154 and 162 of SEQ ID NO: 2 or between amino acid numbers 107 and 115 of SEQ ID NO: 4.

[0021] In still another embodiment, the additional protease recognition site is inserted between amino acids corresponding to amino acid numbers 154 and 160 of SEQ ID NO: 2. Preferably, the additional protease recognition site is inserted between amino acid numbers 154 and 160 of SEQ ID NO: 2 or between amino acid numbers 107 and 113 of SEQ ID NO: 4.

[0022] In a further embodiment, the additional protease recognition site is inserted in domain I between amino acids corresponding to amino acid numbers 154 and 158 of SEQ ID NO: 2. Preferably, the additional protease recognition site is inserted in domain I between amino acid numbers 154 and 158 of SEQ ID NO: 2 or between amino acid numbers 107 and 111 of SEQ ID NO: 4.

[0023] In another embodiment, the additional protease recognition site is inserted in domain III between amino acids corresponding to amino acid numbers 583 and 589 of SEQ ID NO: 2. Preferably, the additional protease site is inserted in domain III between amino acid numbers 583 and 589 of SEQ ID NO: 2 or between amino acid numbers 536 and 542 of SEQ ID NO: 4.

[0024] In still another embodiment, the additional protease recognition site is inserted in domain III between amino acids corresponding to amino acid numbers 583 and 588 of SEQ ID NO: 2. Preferably, the additional protease site is inserted in domain III between amino acid numbers 583 and 588 of SEQ ID NO: 2 or between amino acid numbers 536 and 541of SEQ ID NO: 4.

[0025] In yet another embodiment, the additional protease recognition site is inserted in domain III between amino acids corresponding to amino acid numbers 587 and 588 of SEQ ID NO: 2. Preferably, the additional protease site is inserted in domain III between amino acid numbers 587 and 588 of SEQ ID NO: 2 or between amino acid numbers 540 and 541 of SEQ ID NO: 4.

[0026] In one embodiment, the additional protease recognition site is inserted in domain I and domain III of the unmodified Cry3A toxin. Preferably, the additional protease recognition site is inserted in domain I at a position that replaces or is adjacent to a

60215021.082101

296
298

naturally occurring protease recognition site and in domain III at a position that is within, replaces, or is adjacent to a naturally occurring protease recognition site.

[0027] In another embodiment, the additional protease recognition site is inserted in domain I between amino acids corresponding to amino acid numbers 154 and 160 and in domain III between amino acids corresponding to amino acid numbers 587 and 588 of SEQ ID NO: 2. Preferably, the additional protease recognition site is inserted in domain I between amino acid numbers 154 and 160 and in domain III between amino acid numbers 587 and 588 of SEQ ID NO: 2 or in domain I between amino acid numbers 107 and 113 and in domain III between amino acid numbers 540 and 541 of SEQ ID NO: 4.

[0028] In yet another embodiment, the additional protease recognition site is located in domain I between amino acids corresponding to amino acid numbers 154 and 158 and in domain III between amino acids corresponding to amino acid numbers 587 and 588 of SEQ ID NO: 2. Preferably, the additional protease recognition site is inserted in domain I between amino acid numbers 154 and 158 and in domain III between amino acid numbers 587 and 588 of SEQ ID NO: 2 or in domain I between amino acid numbers 107 and 111 and in domain III between amino acid numbers 540 and 541 of SEQ ID NO: 4.

[0029] In another embodiment, the additional protease recognition site is located in domain I between amino acids corresponding to amino acid numbers 154 and 158 and in domain III between amino acids corresponding to amino acid numbers 583 and 588 of SEQ ID NO: 2. Preferably, the additional protease recognition site is inserted in domain I between amino acid numbers 154 and 158 and in domain III between amino acid numbers 583 and 588 of SEQ ID NO: 2 or in domain I between amino acid numbers 107 and 111 and in domain III between amino acid numbers 536 and 541 of SEQ ID NO: 4.

[0030] In a preferred embodiment, the isolated nucleic acid molecule of the present invention comprises nucleotides 1-1791of SEQ ID NO: 6, nucleotides 1-1806 of SEQ ID NO: 8, nucleotides 1-1818 of SEQ ID NO: 10, nucleotides 1-1794 of SEQ ID NO: 12, nucleotides 1-1812 of SEQ ID NO: 14, nucleotides 1-1812 of SEQ ID NO: 16, nucleotides 1-1818 of SEQ ID NO: 18, or nucleotides 1-1791 of SEQ ID NO: 20.

[0031] In another preferred embodiment, the isolated nucleic acid molecule of the invention encodes a modified Cry3A toxin comprising the amino acid sequence set forth in SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, or SEQ ID NO: 21.

S0316#21.082101

[0032] According to one embodiment of the invention, the isolated nucleic acid molecule encodes a modified Cry3A toxin that is active against a coleopteran insect. Preferably, the modified Cry3A toxin has activity against western corn rootworm.

[0033] The present invention provides a chimeric gene comprising a heterologous promoter sequence operatively linked to the nucleic acid molecule of the invention. The present invention also provides a recombinant vector comprising such a chimeric gene. Further, the present invention provides a transgenic non-human host cell comprising such a chimeric gene. A transgenic host cell according to this aspect of the invention may be a bacterial cell or a plant cell, preferably, a plant cell. The present invention further provides a transgenic plant comprising such a plant cell. A transgenic plant according to this aspect of the invention may be sorghum, wheat, sunflower, tomato, potato, cole crops, cotton, rice, soybean, sugar beet, sugarcane, tobacco, barley, oilseed rape, or maize, preferably, maize. The present invention also provides seed from the group of transgenic plants consisting of sorghum, wheat, sunflower, tomato, potato, cole crops, cotton, rice, soybean, sugar beet, sugarcane, tobacco, barley, oilseed rape, and maize. In a particularly preferred embodiment, the seed is from a transgenic maize plant.

[0034] In another aspect, the present invention provides toxins produced by the expression of the nucleic acid molecules of the present invention. In a preferred embodiment, the toxin is produced by the expression of the nucleic acid molecule comprising nucleotides 1-1791of SEQ ID NO: 6, nucleotides 1-1806 of SEQ ID NO: 8, nucleotides 1-1818 of SEQ ID NO: 10, nucleotides 1-1794 of SEQ ID NO: 12, nucleotides 1-1812 of SEQ ID NO: 14, nucleotides 1-1812 of SEQ ID NO: 16, nucleotides 1-1818 of SEQ ID NO: 18, or nucleotides 1-1791 of SEQ ID NO: 20.

[0035] In another embodiment, the toxins of the invention are active against coleopteran insects, preferably against western corn rootworm.

[0036] In one embodiment, a toxin of the present invention comprises the amino acid sequence set forth in SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, or SEQ ID NO: 21.

[0037] The present invention also provides a composition comprising an effective insect-controlling amount of a toxin according to the invention.

[0038] In another aspect, the present invention provides a method of producing a toxin that is active against insects, comprising: (a) obtaining a host cell comprising a chimeric gene, which itself comprises a heterologous promoter sequence operatively linked to the

SO31GH21.080101

298
300

nucleic acid molecule of the invention; and (b) expressing the nucleic acid molecule in the transgenic host cell, which results in at least one toxin that is active against insects.

[0039] In a further aspect, the present invention provides a method of producing an insect-resistant transgenic plant, comprising introducing a nucleic acid molecule of the invention into the transgenic plant, wherein the nucleic acid molecule is expressible in the transgenic plant in an effective amount to control insects. In a preferred embodiment, the insects are coleopteran insects, preferably western corn rootworm.

[0040] In yet a further aspect, the present invention provides a method of controlling insects, comprising delivering to the insects an effective amount of a toxin of the invention. According to one embodiment, the insects are coleopteran insects, preferably, western corn rootworm.

[0041] Preferably, the toxin is delivered to the insects orally. In one preferred embodiment, the toxin is delivered orally through a transgenic plant comprising a nucleic acid sequence that expresses a toxin of the present invention.

[0042] Also provided by the present invention is a method of making a modified Cry3A toxin, comprising: (a) obtaining a *cry*3A toxin gene which encodes a Cry3A toxin; (b) identifying a gut protease of a target insect; (c) obtaining a nucleotide sequence which encodes a recognition sequence for the gut protease; (d) inserting the nucleotide sequence of (c) into either domain I or domain III or both domain I and domain III at a position that replaces, is within, or adjacent to a nucleotide sequence that codes for a naturally occurring protease recognition site in a *cry*3A toxin gene, thus creating a modified *cry*3A toxin gene; (e) inserting the modified *cry*3A toxin gene in an expression cassette; (f) expressing the modified *cry*3A toxin gene in a non-human host cell, resulting in the host cell producing a modified Cry3A toxin; and, (g) bioassaying the modified Cry3A toxin against a target insect, whereby the modified Cry3A toxin causes higher mortality to the target insect than the mortality caused by a Cry3A toxin. In a preferred embodiment, the modified Cry3A toxin causes at least about 50% mortality to the target insect when the Cry3A toxin causes up to about 30% mortality.

[0043] The present invention further provides a method of controlling insects wherein the transgenic plant further comprises a second nucleic acid sequence or groups of nucleic acid sequences that encode a second pesticidal principle. Particularly preferred second nucleic acid sequences are those that encode a δ-endotoxin, those that encode a Vegetative Insecticidal Protein toxin, disclosed in U.S. Patents 5,849,870 and 5,877,012,

60315421.083101

301

incorporated herein by reference, or those that encode a pathway for the production of a non-proteinaceous pesticidal principle.

[0044] Yet another aspect of the present invention is the provision of a method for mutagenizing a nucleic acid molecule according to the present invention, wherein the nucleic acid molecule has been cleaved into populations of double-stranded random fragments of a desired size, comprising: (a) adding to the population of double-stranded random fragments one or more single- or double-stranded oligonucleotides, wherein the oligonucleotides each comprise an area of identity and an area of heterology to a double-stranded template polynucleotide; (b) denaturing the resultant mixture of double-stranded random fragments and oligonucleotides into single-stranded fragments; (c) incubating the resultant population of single-stranded fragments with polymerase under conditions which result in the annealing of the single-stranded fragments at the areas of identity to form pairs of annealed fragments, the areas of identity being sufficient for one member of the pair to prime replication of the other, thereby forming a mutagenized double-stranded polynucleotide; and (d) repeating the second and third steps for at least two further cycles, wherein the resultant mixture in the second step of a further cycle includes the mutagenized double-stranded polynucleotide from the third step of the previous cycle, and wherein the further cycle forms a further mutagenized double-stranded polynucleotide.

[0045] Other aspects and advantages of the present invention will become apparent to those skilled in the art from a study of the following description of the invention and non-limiting examples.

BRIEF DESCRIPTION OF THE SEQUENCES IN THE SEQUENCE LISTING

SEQ ID NO: 1 is the native *cry*3A coding region.

SEQ ID NO: 2 is the amino acid sequence of the Cry3A toxin encoded by the native *cry*3A gene.

SEQ ID NO: 3 is the maize optimized *cry*3A coding region beginning at nucleotide 144 of the native *cry*3A coding region.

SEQ ID NO: 4 is the amino acid sequence of the Cry3A toxin encoded by the maize optimized *cry*3A gene.

SEQ ID NO: 5 is the nucleotide sequence of pCIB6850.

300
302

SEQ ID NO: 6 is the maize optimized modified *cry*3A054 coding sequence.

SEQ ID NO: 7 is the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 6.

SEQ ID NO: 8 is the maize optimized modified *cry*3A055 coding sequence.

SEQ ID NO: 9 is the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 8.

SEQ ID NO: 10 is the maize optimized modified *cry*3A085 coding sequence.

SEQ ID NO: 11 is the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 10.

SEQ ID NO: 12 is the maize optimized modified *cry*3A082 coding sequence.

SEQ ID NO: 13 is the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 12.

SEQ ID NO: 14 is the maize optimized modified *cry*3A058 coding sequence.

SEQ ID NO: 15 is the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 14.

SEQ ID NO: 16 is the maize optimized modified *cry*3A057 coding sequence.

SEQ ID NO: 17 is the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 16.

SEQ ID NO: 18 is the maize optimized modified *cry*3A056 coding sequence.

SEQ ID NO: 19 is the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 18.

SEQ ID NO: 20 is the maize optimized modified *cry*3A083 coding sequence.

SEQ ID NO: 21 is the amino acid sequence encoded by the nucleotide sequence of SEQ ID NO: 20.

SEQ ID NOS: 22-34 are PCR primers useful in the present invention.

SEQ ID NO: 35 is an amino acid sequence comprising a cathepsin G recognition site.

SEQ ID NO: 36 is an amino acid sequence comprising a cathepsin G recognition site.

SEQ ID NO: 37 is an amino acid sequence comprising a cathepsin G recognition site.

SEQ ID NO: 38 is an amino acid sequence comprising a cathepsin G recognition site.

60316421.082101

DEFINITIONS

301
303

[0046] For clarity, certain terms used in the specification are defined and presented as follows:

[0047] "Activity" of the modified Cry3A toxins of the invention is meant that the modified Cry3A toxins function as orally active insect control agents, have a toxic effect, or are able to disrupt or deter insect feeding, which may or may not cause death of the insect. When a modified Cry3A toxin of the invention is delivered to the insect, the result is typically death of the insect, or the insect does not feed upon the source that makes the modified Cry3A toxin available to the insect.

[0048] "Adjacent to"- According to the present invention, an additional protease recognition site is "adjacent to" a naturally occurring protease recognition site when the additional protease recognition site is within four residues, preferably within three residues, more preferably within two residues, and most preferably within one residue of a naturally occurring protease recognition site. For example, an additional protease recognition site inserted between Pro-154 and Arg-158 of the deduced amino acid sequence of a Cry3A toxin (SEQ ID NO: 2) is "adjacent to" the naturally occurring trypsin recognition site located between Arg-158 and Asn-159 of the deduced amino acid sequence of the Cry3A toxin (SEQ ID NO: 2).

[0049] The phrase "approximately the same position" as used herein to describe the location where an additional protease recognition site is inserted into a Cry3A toxin in relation to a naturally occurring protease recognition site, means that the location is at most four residues away from a naturally occurring protease recognition site. The location can also be three or two residues away from a naturally occurring protease recognition site. The location can also be one residue away from a naturally occurring protease recognition site. "Approximately the same position" can also mean that the additional protease recognition site is inserted within a naturally occurring protease recognition site.

[0050] "Associated with / operatively linked" refer to two nucleic acid sequences that are related physically or functionally. For example, a promoter or regulatory DNA sequence is said to be "associated with" a DNA sequence that codes for an RNA or a protein if the two sequences are operatively linked, or situated such that the regulatory DNA sequence will affect the expression level of the coding or structural DNA sequence.

[0051] A "chimeric gene" or "chimeric construct" is a recombinant nucleic acid sequence in which a promoter or regulatory nucleic acid sequence is operatively linked to, or

S0215pct1 - 83101

302
304

associated with, a nucleic acid sequence that codes for an mRNA or which is expressed as a protein, such that the regulatory nucleic acid sequence is able to regulate transcription or expression of the associated nucleic acid coding sequence. The regulatory nucleic acid sequence of the chimeric gene is not normally operatively linked to the associated nucleic acid sequence as found in nature.

[0052] A "coding sequence" is a nucleic acid sequence that is transcribed into RNA such as mRNA, rRNA, tRNA, snRNA, sense RNA or antisense RNA. Preferably the RNA is then translated in an organism to produce a protein.

[0053] To "control" insects means to inhibit, through a toxic effect, the ability of insect pests to survive, grow, feed, and/or reproduce, or to limit insect-related damage or loss in crop plants. To "control" insects may or may not mean killing the insects, although it preferably means killing the insects.

[0054] Corresponding to: in the context of the present invention, "corresponding to" means that when the amino acid sequences of variant Cry3A δ-endotoxins are aligned with each other, the amino acids that "correspond to" certain enumerated positions in the present invention are those that align with these positions in the Cry3A toxin (SEQ ID NO: 2), but that are not necessarily in these exact numerical positions relative to the particular Cry3A amino acid sequence of the invention. For example, the maize optimized *cry*3A gene (SEQ ID NO: 3) of the invention encodes a Cry3A toxin (SEQ ID NO: 4) that begins at *Met*-48 of the Cry3A toxin (SEQ ID NO: 2) encoded by the native *cry*3A gene (SEQ ID NO: 1). Therefore, according to the present invention, amino acid numbers 107-115, including all numbers in between, and 536-541, including all numbers in between, of SEQ ID NO: 4 correspond to amino acid numbers 154-163, and all numbers in between, and 583-588, and all numbers in between, respectively, of SEQ ID NO: 2.

[0055] A "Cry3A toxin", as used herein, refers to an approximately 73 kDa *Bacillus thuringiensis* var. *tenebrionis* (Krieg *et al.*, 1983, Z. Angew. Entomol. 96:500-508) (Bt) coleopteran-active protein (Sekar *et al.*, 1987, Proc. Nalt. Acad. Sci. 84:7036-7040), for example SEQ ID NO: 2, as well as any truncated lower molecular weight variants, derivable from a Cry3A toxin, for example SEQ ID NO: 4, and retaining substantially the same toxicity as the Cry3A toxin. The lower molecular weight variants can be obtained by protease cleavage of naturally occurring protease recognition sites of the Cry3A toxin or by a second translational initiation codon in the same frame as the translational

305
305

initiation codon coding for the 73 kDa Cry3A toxin. The amino acid sequence of a Cry3A toxin and the lower molecular weight variants thereof can be found in a toxin naturally occurring in Bt. A Cry3A toxin can be encoded by a native Bt gene as in SEQ ID NO: 1 or by a synthetic coding sequence as in SEQ ID NO: 3. A "Cry3A toxin" does not have any additional protease recognition sites over the protease recognition sites that naturally occur in the Cry3A toxin. A Cry3A toxin can be isolated, purified or expressed in a heterologous system.

[0056] A "*cry*3A gene", as used herein, refers to the nucleotide sequence of SEQ ID NO: 1 or SEQ ID NO: 3. A *cry*3A gene (Sekar *et al.*, 1987, Proc. Natl. Acad. Sci. 84:7036-7040) can be naturally occurring, as found in *Bacillus thuringiensis* var. *tenebrionis* (Kreig *et al.*, 1983, Z. Angew. Entomol. 96:500-508), or synthetic and encodes a Cry3A toxin. The *cry*3A gene of this invention can be referred to as the native *cry*3A gene as in SEQ ID NO: 1 or the maize-optimized *cry*3A gene as in SEQ ID NO: 3.

[0057] To "deliver" a toxin means that the toxin comes in contact with an insect, resulting in toxic effect and control of the insect. The toxin can be delivered in many recognized ways, e.g., orally by ingestion by the insect or by contact with the insect via transgenic plant expression, formulated protein composition(s), sprayable protein composition(s), a bait matrix, or any other art-recognized toxin delivery system.

[0058] "Effective insect-controlling amount" means that concentration of toxin that inhibits, through a toxic effect, the ability of insects to survive, grow, feed and/or reproduce, or to limit insect-related damage or loss in crop plants. "Effective insect-controlling amount" may or may not mean killing the insects, although it preferably means killing the insects.

[0059] "Expression cassette" as used herein means a nucleic acid sequence capable of directing expression of a particular nucleotide sequence in an appropriate host cell, comprising a promoter operably linked to the nucleotide sequence of interest which is operably linked to termination signals. It also typically comprises sequences required for proper translation of the nucleotide sequence. 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 that 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 nucleic acid sequence

304
306

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 that 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.

[0060] A "gene" is a defined region that is located within a genome and that, besides the aforementioned coding nucleic acid sequence, comprises other, primarily regulatory, nucleic acid sequences responsible for the control of the expression, that is to say the transcription and translation, of the coding portion. A gene may also comprise other 5' and 3' untranslated sequences and termination sequences. Further elements that may be present are, for example, introns.

[0061] "Gene of interest" refers to any gene which, when transferred to a plant, confers upon the plant a desired characteristic such as antibiotic resistance, virus resistance, insect resistance, disease resistance, or resistance to other pests, herbicide tolerance, improved nutritional value, improved performance in an industrial process or altered reproductive capability. The "gene of interest" may also be one that is transferred to plants for the production of commercially valuable enzymes or metabolites in the plant.

[0062] A "gut protease" is a protease naturally found in the digestive tract of an insect. This protease is usually involved in the digestion of ingested proteins.

[0063] A "heterologous" nucleic acid sequence is a nucleic acid sequence not naturally associated with a host cell into which it is introduced, including non- naturally occurring multiple copies of a naturally occurring nucleic acid sequence.

[0064] A "homologous" nucleic acid sequence is a nucleic acid sequence naturally associated with a host cell into which it is introduced.

[0065] "Homologous recombination" is the reciprocal exchange of nucleic acid fragments between homologous nucleic acid molecules.

[0066] "Insecticidal" is defined as a toxic biological activity capable of controlling insects, preferably by killing them.

[0067] A nucleic acid sequence is "isocoding with" a reference nucleic acid sequence when the nucleic acid sequence encodes a polypeptide having the same amino acid sequence as the polypeptide encoded by the reference nucleic acid sequence.

305
307

[0068] An "isolated" nucleic acid molecule or an isolated toxin is a nucleic acid molecule or toxin that, by the hand of man, exists apart from its native environment and is therefore not a product of nature. An isolated nucleic acid molecule or toxin may exist in a purified form or may exist in a non-native environment such as, for example, a recombinant host cell.

[0069] A "modified Cry3A toxin" of this invention, refers to a Cry3A-derived toxin having at least one additional protease recognition site that is recognized by a gut protease of a target insect, which does not naturally occur in a Cry3A toxin. A modified Cry3A toxin is not naturally occurring and, by the hand of man, comprises an amino acid sequence that is not identical to a naturally occurring toxin found in *Bacillus thuringiensis*. The modified Cry3A toxin causes higher mortality to a target insect than the mortality caused by a Cry3A toxin to the same target insect.

[0070] A "modified *cry*3A gene" according to this invention, refers to a *cry*3A-derived gene comprising the coding sequence of at least one additional protease recognition site that does not naturally occur in an unmodified *cry*3A gene. The modified *cry*3A gene can be derived from a native *cry*3A gene or from a synthetic *cry*3A gene.

[0071] A "naturally occurring protease recognition site" is a location within a Cry3A toxin that is cleaved by a non-insect derived protease or by a protease or gut extract from an insect species susceptible to the Cry3A toxin. For example, a naturally occurring protease recognition site, recognized by trypsin and proteases found in a susceptible insect gut extract, exists between Arg-158 and Asn-159 of the deduced Cry3A toxin amino acid sequence (SEQ ID NO: 2). Naturally occurring protease recognition sites, recognized by chymotrypsin, exist between His-161 and Ser-162 as well as between Tyr-587 and Tyr-588 of the deduced Cry3A toxin amino acid sequence (SEQ ID NO: 2).

[0072] A "nucleic acid molecule" or "nucleic acid sequence" is a linear segment of single- or double-stranded DNA or RNA that can be isolated from any source. In the context of the present invention, the nucleic acid molecule is preferably a segment of DNA.

[0073] A "plant" is any plant at any stage of development, particularly a seed plant.

[0074] A "plant cell" is a structural and physiological unit of a plant, comprising a protoplast and a cell wall. The plant cell may be in the form of an isolated single cell or a cultured cell, or as a part of a higher organized unit such as, for example, plant tissue, a plant organ, or a whole plant.

306
308

[0075] "Plant cell culture" means cultures of plant units such as, for example, protoplasts, cell culture cells, cells in plant tissues, pollen, pollen tubes, ovules, embryo sacs, zygotes and embryos at various stages of development.

[0076] "Plant material" refers to leaves, stems, roots, flowers or flower parts, fruits, pollen, egg cells, zygotes, seeds, cuttings, cell or tissue cultures, or any other part or product of a plant.

[0077] A "plant organ" is a distinct and visibly structured and differentiated part of a plant such as a root, stem, leaf, flower bud, or embryo.

[0078] "Plant tissue" as used herein means a group of plant cells organized into a structural and functional unit. Any tissue of a plant *in planta* or in culture is included. This term includes, but is not limited to, whole plants, plant organs, plant seeds, tissue culture and any groups of plant cells organized into structural and/or functional units. The use of this term in conjunction with, or in the absence of, any specific type of plant tissue as listed above or otherwise embraced by this definition is not intended to be exclusive of any other type of plant tissue.

[0079] A "promoter" is an untranslated DNA sequence upstream of the coding region that contains the binding site for RNA polymerase and initiates transcription of the DNA. The promoter region may also include other elements that act as regulators of gene expression.

[0080] A "protoplast" is an isolated plant cell without a cell wall or with only parts of the cell wall.

[0081] "Regulatory elements" refer to sequences involved in controlling the expression of a nucleotide sequence. Regulatory elements comprise a promoter operably linked to the nucleotide sequence of interest and termination signals. They also typically encompass sequences required for proper translation of the nucleotide sequence.

[0082] "Replaces" a naturally occurring protease recognition site - According to the present invention, an additional protease recognition site "replaces" a naturally occurring protease recognition site when insertion of the additional protease recognition site eliminates the naturally occurring protease recognition site. For example, an additional protease recognition site inserted between Pro-154 and Pro-160 of the deduced amino acid sequence of a Cry3A toxin (SEQ ID NO: 2) which eliminates the Arg-158 and Asn-159 residues "replaces" the naturally occurring trypsin recognition site located between Arg-158 and Asn-159 of the deduced amino acid sequence of the Cry3A toxin (SEQ ID NO: 2).

307
309

[0083] "Serine proteases". describe the same group of enzymes that catalyze the hydrolysis of covalent peptidic bonds using a mechanism based on nucleophilic attack of the targeted peptidic bond by a serine. Serine proteases are sequence specific. That is, each serine protease recognizes a specific sub-sequence within a protein where enzymatic recognition occurs.

[0084] A "target insect" is an insect pest species that has little or no susceptibility to a Cry3A toxin and is identified as being a candidate for using the technology of the present invention to control. This control can be achieved through several means but most preferably through the expression of the nucleic acid molecules of the invention in transgenic plants.

[0085] A "target insect gut protease" is a protease found in the gut of a target insect whose recognition site can be inserted into a Cry3A toxin to create a modified Cry3A toxin of the invention.

[0086] "Transformation" is a process for introducing heterologous nucleic acid into a host cell or organism. In particular, "transformation" means the stable integration of a DNA molecule into the genome of an organism of interest.

[0087] "Transformed / transgenic / recombinant" refer to a host organism such as a bacterium or a plant into which a heterologous nucleic acid molecule has been introduced. The nucleic acid molecule can be stably integrated into the genome of the host or the nucleic acid molecule can also be present as an extrachromosomal molecule. Such an extrachromosomal molecule can be auto-replicating. Transformed cells, tissues, or plants are understood to encompass not only the end product of a transformation process, but also transgenic progeny thereof. A "non-transformed", "non-transgenic", or "non-recombinant" host refers to a wild-type organism, e.g., a bacterium or plant, which does not contain the heterologous nucleic acid molecule.

[0088] "Within" a naturally occurring protease recognition site – According to the present invention, an additional protease recognition site is "within" a naturally occurring protease recognition site when the additional protease recognition site lies between the amino acid residue that comes before and the amino acid residue that comes after the naturally occurring protease recognition site. For example, an additional protease recognition site inserted between Tyr-587 and Tyr-588 of the deduced amino acid sequence of a Cry3A toxin (SEQ ID NO: 2) is "within" a naturally occurring chymotrypain recognition site located between Tyr-587 and Tyr-588 of the deduced

amino acid sequence of the Cry3A toxin (SEQ ID NO: 2). The insertion of an additional protease recognition site within a naturally occurring protease recognition site may or may not change the recognition of the naturally occurring protease recognition site by a protease.

[0089] Nucleotides are indicated by their bases by the following standard abbreviations: adenine (A), cytosine (C), thymine (T), and guanine (G). Amino acids are likewise indicated by the following standard abbreviations: alanine (Ala; A), arginine (Arg; R), asparagine (Asn; N), aspartic acid (Asp; D), cysteine (Cys; C), glutamine (Gln; Q), glutamic acid (Glu; E), glycine (Gly; G), histidine (His; H), isoleucine (Ile; I), leucine (Leu; L), lysine (Lys; K), methionine (Met; M), phenylalanine (Phe; F), proline (Pro; P), serine (Ser; S), threonine (Thr; T), tryptophan (Trp; W), tyrosine (Tyr; Y), and valine (Val; V).

DESCRIPTION

[0090] This invention relates to modified *cry*3A nucleic acid sequences whose expression results in modified Cry3A toxins, and to the making and using of the modified Cry3A toxins to control insect pests. The expression of the modified *cry*3A nucleic acid sequences results in modified Cry3A toxins that can be used to control coleopteran insects such as western corn rootworm and northern corn rootworm. A modified Cry3A toxin of the present invention comprises at least one additional protease recognition site that does not naturally occur in a Cry3A toxin. The additional protease recognition site, which is recognized by a gut protease of a target insect, is inserted at approximately the same position as a naturally occurring protease recognition site in a Cry3A toxin. The modified Cry3A toxin causes higher mortality to a target insect than the mortality caused by a Cry3A toxin to the same target insect. Preferably, the modified Cry3A toxin causes at least about 50 % mortality to the target insect to which a Cry3A toxin causes up to about 30% mortality.

[0091] In one preferred embodiment, the invention encompasses an isolated nucleic acid molecule that encodes a modified Cry3A toxin, wherein the additional protease recognition site is recognized by the target insect gut protease, cathepsin G. Cathepsin G activity is determined to be present in the gut of the target insect, western corn rootworm, as described in Example 2. Preferably, the substrate amino acid sequence, AAPF (SEQ ID

NO: 35), used to determine the presence of the cathepsin G activity is inserted into the Cry3A toxin according to the present invention. Other cathepsin G recognition sites can also be used according to the present invention, for example, AAPM (SEQ ID NO: 36), AVPF (SEQ ID NO: 37), PFLF (SEQ ID NO: 38) or other cathepsin G recognition sites as determined by the method of Tanaka *et al.*, 1985 (Biochemistry 24:2040-2047), incorporated herein by reference. Protease recognition sites of other proteases identified in a target insect gut can be used, for example, protease recognition sites recognized by other serine proteases, cysteine proteases and aspartic proteases. Preferable serine proteases encompassed by this embodiment include trypsin, chymotrypsin, carboxypeptidase, endopeptidase and elastase.

[0092] In another preferred embodiment, the invention encompasses an isolated nucleic acid molecule that encodes a modified Cry3A toxin wherein the additional protease recognition site is inserted in either domain I or domain III or in both domain I and domain III of the Cry3A toxin. Preferably, the additional protease recognition site is inserted in domain I, domain III, or domain I and domain III at a position that replaces, is adjacent to, or is within a naturally occurring protease recognition site in the Cry3A toxin. Specifically exemplified herein are nucleic acid molecules that encode modified Cry3A toxins that comprise a cathepsin G recognition site inserted in domain I, domain III, or domain I and domain III at a position that replaces, is adjacent to, or is within a naturally occurring protease recognition site in the unmodified Cry3A toxin.

[0093] Specifically exemplified teachings of methods to make modified *cry*3A nucleic acid molecules that encode modified Cry3A toxins can be found in Example 3. Those skilled in the art will recognize that other methods known in the art can also be used to insert additional protease recognition sites into Cry3A toxins according to the present invention.

[0094] In another preferred embodiment, the invention encompasses an isolated nucleic acid molecule that encodes a modified Cry3A toxin wherein the additional protease recognition site is inserted in domain I between amino acids corresponding to amino acid numbers 154 and 162 of SEQ ID NO: 2. Preferably, the additional protease recognition site is inserted between amino acid numbers 154 and 162 of SEQ ID NO: 2 or between amino acid numbers 107 and 115 of SEQ ID NO: 4. In a preferred embodiment, the additional protease recognition site is inserted between amino acids corresponding to amino acid numbers 154 and 160 of SEQ ID NO: 2. Preferably, the additional protease

recognition site is inserted between amino acid number 154 and 160 of SEQ ID NO: 2 or between amino acid numbers 107 and 113 of SEQ ID NO: 4. Specifically exemplified herein is a nucleic acid molecule, designated *cry*3A054 (SEQ ID NO: 6), that encodes the modified Cry3A054 toxin (SEQ ID NO: 7) comprising a cathepsin G recognition site inserted in domain I between amino acid numbers 107 and 113 of SEQ ID NO: 4. The cathepsin G recognition site replaces a naturally occurring trypsin recognition site and is adjacent to a naturally occurring chymotrypsin recognition site. When expressed in a heterologous host, the nucleic acid molecule of SEQ ID NO: 6 results in insect control activity against western corn rootworm and northern corn rootworm, showing that the nucleic acid sequence set forth in SEQ ID NO: 6 is sufficient for such insect control activity.

[0095] In another preferred embodiment, the additional protease recognition site is inserted in domain I between amino acids corresponding to amino acid numbers 154 and 158 of SEQ ID NO: 2. Preferably, the additional protease recognition site is inserted in domain I between amino acid numbers 154 and 158 of SEQ ID NO: 2 or between amino acid numbers 107 and 111 of SEQ ID NO: 4. Specifically exemplified herein are nucleic acid molecules, designated *cry*3A055 (SEQ ID NO: 8), that encodes the modified Cry3A055 toxin (SEQ ID NO: 9), and *cry*3A085 (SEQ ID NO: 10), that encodes the modified Cry3A085 toxin (SEQ ID NO: 11), comprising a cathepsin G recognition site inserted in domain I between amino acid numbers 107 and 111 of SEQ ID NO: 4. The cathepsin G recognition site is adjacent to naturally occurring trypsin and chymotrypsin recognition sites. When expressed in a heterologous host, the nucleic acid molecule of SEQ ID NO: 8 or SEQ ID NO: 10 results in insect control activity against western corn rootworm and northern corn rootworm, showing that the nucleic acid sequence set forth in SEQ ID NO: 8 or SEQ ID NO: 10 is sufficient for such insect control activity.

[0096] In a preferred embodiment, the invention encompasses an isolated nucleic acid molecule that encodes a modified Cry3A toxin wherein the additional protease recognition site is inserted in domain III between amino acids corresponding to amino acid numbers 583 and 589 of SEQ ID NO: 2. Preferably, the additional protease site is inserted in domain III between amino acid numbers 583 and 589 of SEQ ID NO: 2 or between amino acid numbers 536 and 542 of SEQ ID NO: 4.

[0097] In another preferred embodiment, the invention encompasses an isolated nucleic acid molecule that encodes a modified Cry3A toxin wherein the additional protease

recognition site is inserted in domain III between amino acids corresponding to amino acid numbers 583 and 588 of SEQ ID NO: 2. Preferably, the additional protease site is inserted in domain III between amino acid numbers 583 and 588 of SEQ ID NO: 2 or between amino acid numbers 536 and 541 of SEQ ID NO: 4. Specifically exemplified herein is a nucleic acid molecule, designated *cry*3A082 (SEQ ID NO: 12), that encodes the modified Cry3A082 toxin (SEQ ID NO: 13) comprising a cathepsin G recognition site inserted in domain III between amino acid numbers 536 and 541 of SEQ ID NO: 4. The cathepsin G recognition site replaces a naturally occurring chymotrypsin recognition site. When expressed in a heterologous host, the nucleic acid molecule of SEQ ID NO: 12 results in insect control activity against western corn rootworm and northern corn rootworm, showing that the nucleic acid sequence set forth in SEQ ID NO: 12 is sufficient for such insect control activity.

[0098] In another preferred embodiment, the additional protease recognition site is inserted in domain III between amino acids corresponding to amino acid numbers 587 and 588 of SEQ ID NO: 2. Preferably, the additional protease site is inserted in domain III between amino acid numbers 587 and 588 of SEQ ID NO: 2 or between amino acid numbers 540 and 541 of SEQ ID NO: 4. Specifically exemplified herein is a nucleic acid molecule, designated *cry*3A058 (SEQ ID NO: 14), that encodes the modified Cry3A058 toxin (SEQ ID NO: 15) comprising a cathepsin G recognition site inserted in domain III between amino acid numbers 540 and 541 of SEQ ID NO: 4. The cathepsin G recognition site is within a naturally occurring chymotrypsin recognition site. When expressed in a heterologous host, the nucleic acid molecule of SEQ ID NO: 14 results in insect control activity against western corn rootworm and northern corn rootworm, showing that the nucleic acid sequence set forth in SEQ ID NO: 14 is sufficient for such insect control activity.

[0099] In yet another preferred embodiment, the invention encompasses an isolated nucleic acid molecule that encodes a modified Cry3A toxin wherein the additional protease recognition site is inserted in domain I between amino acids corresponding to amino acid numbers 154 and 160 and in domain III between amino acids corresponding to amino acid numbers 587 and 588 of SEQ ID NO: 2. Preferably, the additional protease recognition site is inserted in domain I between amino acid numbers 154 and 160 and in domain III between amino acid numbers 587 and 588 of SEQ ID NO: 2 or in domain I between amino acid numbers 107 and 113 and in domain III between amino acid numbers

312
314

540 and 541 of SEQ ID NO: 4. Specifically exemplified herein is a nucleic acid molecule, designated *cry*3A057 (SEQ ID N : 16), that encodes the modified Cry3A057 toxin (SEQ ID NO: 17) comprising a cathepsin G recognition site inserted in domain I between amino acid numbers 107 and 113 and in domain III between amino acid numbers 540 and 541 of SEQ ID NO: 4. The cathepsin G recognition site replaces a naturally occurring trypsin recognition site and is adjacent to a naturally occurring chymotrypsin recognition site in domain I and is within a naturally occurring chymotrypsin recognition site in domain III. When expressed in a heterologous host, the nucleic acid molecule of SEQ ID NO: 16 results in insect control activity against western corn rootworm and northern corn rootworm, showing that the nucleic acid sequence set forth in SEQ ID NO: 16 is sufficient for such insect control activity.

[00100] In yet another preferred embodiment, the additional protease recognition site is located in domain I between amino acids corresponding to amino acid numbers 154 and 158 and in domain III between amino acids corresponding to amino acid numbers 587 and 588 of SEQ ID NO: 2. Preferably, the additional protease recognition site is inserted in domain I between amino acid numbers 154 and 158 and in domain III between amino acid numbers 587 and 588 of SEQ ID NO: 2 or in domain I between amino acid numbers 107 and 111 and in domain III between amino acid numbers 540 and 541 of SEQ ID NO: 4. Specifically exemplified herein is the nucleic acid molecule designated *cry*3A056 (SEQ ID NO: 18), which encodes the modified Cry3A056 toxin (SEQ ID NO: 19) comprising a cathepsin G recognition site inserted in domain I between amino acid numbers 107 and 111 and in domain III between amino acid numbers 540 and 541 of SEQ ID NO: 4. The cathepsin G recognition site is adjacent to naturally occurring trypsin and chymotrypsin recognition sites in domain I and is within a naturally occurring chymotrypsin recognition site in domain III. When expressed in a heterologous host, the nucleic acid molecule of SEQ ID NO: 18 results in insect control activity against western corn rootworm and northern corn rootworm, showing that the nucleic acid sequence set forth in SEQ ID NO: 18 is sufficient for such insect control activity.

[00101] In still another preferred embodiment, the additional protease recognition site is located in domain I between amino acids corresponding to amino acid numbers 154 and 158 and in domain III between amino acids corresponding to amino acid numbers 583 and 588 of SEQ ID NO: 2. Preferably, the additional protease recognition site is inserted in domain I between amino acid numbers 154 and 158 and in domain III between amino

acid numbers 583 and 588 of SEQ ID NO: 2 or in domain I between amino acid numbers 107 and 111 and in domain III between amino acid numbers 536 and 541 of SEQ ID NO: 4. Specifically exemplified herein is a nucleic acid molecule, designated *cry*3A083 (SEQ ID NO: 20), which encodes the modified Cry3A083 toxin (SEQ ID NO: 21) comprising a cathepsin G recognition site inserted in domain I between amino acid numbers 107 and 111 and in domain III between amino acid numbers 536 and 541 of SEQ ID NO: 4. The cathepsin G recognition site is adjacent to naturally occurring trypsin and chymotrypsin recognition sites in domain I and replaces a naturally occurring chymotrypsin recognition site in domain III. When expressed in a heterologous host, the nucleic acid molecule of SEQ ID NO: 20 results in insect control activity against western corn rootworm and northern corn rootworm, showing that the nucleic acid sequence set forth in SEQ ID NO: 20 is sufficient for such insect control activity.

[00102] In a preferred embodiment, the isolated nucleic acid molecule of the present invention comprises nucleotides 1-1791of SEQ ID NO: 6, nucleotides 1-1806 of SEQ ID NO: 8, nucleotides 1-1812 of SEQ ID NO: 10, nucleotides 1-1794 of SEQ ID NO: 12, nucleotides 1-1818 of SEQ ID NO: 14, nucleotides 1-1812 of SEQ ID NO: 16, nucleotides 1-1791 of SEQ ID NO: 18, and nucleotides 1-1818 of SEQ ID NO: 20.

[00103] In another preferred embodiment, the invention encompasses the isolated nucleic acid molecule that encodes a modified Cry3A toxin comprising the amino acid sequence set forth in SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, or SEQ ID NO: 21.

[00104] The present invention also encompasses recombinant vectors comprising the nucleic acid sequences of this invention. In such vectors, the nucleic acid sequences are preferably comprised in expression cassettes comprising regulatory elements for expression of the nucleotide sequences in a host cell capable of expressing the nucleotides sequences. Such regulatory elements usually comprise promoter and termination signals and preferably also comprise elements allowing efficient translation of polypeptides encoded by the nucleic acid sequences of the present invention. Vectors comprising the nucleic acid sequences are usually capable of replication in particular host cells, preferably as extrachromosomal molecules, and are therefore used to amplify the nucleic acid sequences of this invention in the host cells. In one embodiment, host cells for such vectors are microorganisms, such as bacteria, in particular *Bacillus thuringiensis* or *E. coli*. In another embodiment, host cells for such recombinant vectors are endophytes or

epiphytes. A preferred host cell for such vectors is a eukaryotic cell, such as a plant cell. Plant cells such as maize cells are most preferred host cells. In another preferred embodiment, such vectors are viral vectors and are used for replication of the nucleotide sequences in particular host cells, e.g. insect cells or plant cells. Recombinant vectors are also used for transformation of the nucleotide sequences of this invention into host cells, whereby the nucleotide sequences are stably integrated into the DNA of such host cells. In one, such host cells are prokaryotic cells. In a preferred embodiment, such host cells are eukaryotic cells, such as plant cells. In a most preferred embodiment, the host cells are plant cells, such as maize cells.

[00105] In another aspect, the present invention encompasses modified Cry3A toxins produced by the expression of the nucleic acid molecules of the present invention.

[00106] In preferred embodiments, the modified Cry3A toxins of the invention comprise a polypeptide encoded by a nucleotide sequence of the invention. In a further preferred embodiment, the modified Cry3A toxin is produced by the expression of the nucleic acid molecule comprising nucleotides 1-1791of SEQ ID NO: 6, nucleotides 1-1806 of SEQ ID NO: 8, nucleotides 1-1812 of SEQ ID NO: 10, nucleotides 1-1794 of SEQ ID NO: 12, nucleotides 1-1818 of SEQ ID NO: 14, nucleotides 1-1812 of SEQ ID NO: 16, nucleotides 1-1791 of SEQ ID NO: 18, and nucleotides 1-1818 of SEQ ID NO: 20.

[00107] In a preferred embodiment, a modified Cry3A toxin of the present invention comprises the amino acid sequence set forth in SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, or SEQ ID NO: 21.

[00108] The modified Cry3A toxins of the present invention have insect control activity when tested against insect pests in bioassays. In another preferred embodiment, the modified Cry3A toxins of the invention are active against coleopteran insects, preferably against western corn rootworm and northern corn rootworm. The insect controlling properties of the modified Cry3A toxins of the invention are further illustrated in Examples 4 and 6.

[00109] The present invention also encompasses a composition comprising an effective insect-controlling amount of a modified Cry3A toxin according to the invention.

[00110] In another preferred embodiment, the invention encompasses a method of producing a modified Cry3A toxin that is active against insects, comprising: (a) obtaining a host cell comprising a chimeric gene, which itself comprises a heterologous promoter

sequence operatively linked to the nucleic acid molecule of the invention; and (b) expressing the nucleic acid molecule in the transgenic host cell, which results in at least one modified Cry3A toxin that is active against insects.

[00111] In a further preferred embodiment, the invention encompasses a method of producing an insect-resistant transgenic plant, comprising introducing a nucleic acid molecule of the invention into the transgenic plant, wherein the nucleic acid molecule is expressible in the transgenic plant in an effective amount to control insects. In a preferred embodiment, the insects are coleopteran insects, preferably western corn rootworm and northern corn rootworm.

[00112] In yet a further preferred embodiment, the invention encompasses a method of controlling insects, comprising delivering to the insects an effective amount of a modified Cry3A toxin of the invention. According to this embodiment, the insects are coleopteran insects, preferably, western corn rootworm and northern corn rootworm. Preferably, the modified Cry3A toxin is delivered to the insects orally. In one preferred aspect, the toxin is delivered orally through a transgenic plant comprising a nucleic acid sequence that expresses a modified Cry3A toxin of the present invention.

[00113] The present invention also encompasses a method of making a modified Cry3A toxin, comprising: (a) obtaining a *cry*3A toxin gene which encodes a Cry3A toxin; (b) identifying a gut protease of a target insect; (c) obtaining a nucleotide sequence which encodes a recognition site for the gut protease; (d) inserting the nucleotide sequence of (c) into either domain I or domain III or both domain I and domain III at a position that replaces, is within, or adjacent to a nucleotide sequence that codes for a naturally occurring protease recognition site in the *cry*3A toxin gene, thus creating a modified *cry*3A toxin gene; (e) inserting the modified *cry*3A toxin gene in an expression cassette; (f) expressing the modified *cry*3A toxin gene in a non-human host cell, resulting in the host cell producing a modified Cry3A toxin; and, (g) bioassaying the modified Cry3A toxin against a target insect, which causes higher mortality to the target insect than the mortality caused by a Cry3A toxin. In a preferred embodiment, the modified Cry3A toxin causes at least about 50% mortality to the target insect when the Cry3A toxin causes up to about 30% mortality.

[00114] The present invention further encompasses a method of controlling insects wherein the transgenic plant further comprises a second nucleic acid sequence or groups of nucleic acid sequences that encode a second pesticidal principle. Particularly preferred

second nucleic acid sequences are those that encode a δ-endotoxin, those that encode a Vegetative Insecticidal Protein toxin, disclosed in U.S. Patents 5,849,870 and 5,877,012, incorporated herein by reference, or those that encode a pathway for the production of a non-proteinaceous principle.

[00115] In further embodiments, the nucleotide sequences of the invention can be further modified by incorporation of random mutations in a technique known as *in vitro* recombination or DNA shuffling. This technique is described in Stemmer *et al.*, Nature 370:389-391 (1994) and U.S. Patent 5,605,793, which are incorporated herein by reference. Millions of mutant copies of a nucleotide sequence are produced based on an original nucleotide sequence of this invention and variants with improved properties, such as increased insecticidal activity, enhanced stability, or different specificity or ranges of target-insect pests are recovered. The method encompasses forming a mutagenized double-stranded polynucleotide from a template double-stranded polynucleotide comprising a nucleotide sequence of this invention, wherein the template double-stranded polynucleotide has been cleaved into double-stranded-random fragments of a desired size, and comprises the steps of adding to the resultant population of double-stranded random fragments one or more single or double-stranded oligonucleotides, wherein said oligonucleotides comprise an area of identity and an area of heterology to the double-stranded template polynucleotide; denaturing the resultant mixture of double-stranded random fragments and oligonucleotides into single-stranded fragments; incubating the resultant population of single-stranded fragments with a polymerase under conditions which result in the annealing of said single- stranded fragments at said areas of identity to form pairs of annealed fragments, said areas of identity being sufficient for one member of a pair to prime replication of the other, thereby forming a mutagenized double-stranded polynucleotide; and repeating the second and third steps for at least two further cycles, wherein the resultant mixture in the second step of a further cycle includes the mutagenized double-stranded polynucleotide from the third step of the previous cycle, and the further cycle forms a further mutagenized double-stranded polynucleotide. In a preferred embodiment, the concentration of a single species of double- stranded random fragment in the population of double-stranded random fragments is less than 1% by weight of the total DNA. In a further preferred embodiment, the template double-stranded polynucleotide comprises at least about 100 species of polynucleotides. In another preferred embodiment, the size of the double-stranded random fragments is from about 5

bp to 5 kb. In a further preferred embodiment, the fourth step of the method comprises repeating the second and the third steps for at least 10 cycles.

Expression of the Nucleotide Sequences in Heterologous Microbial Hosts

[00116] As biological insect control agents, the insecticidal modified Cry3A toxins are produced by expression of the nucleotide sequences in heterologous host cells capable of expressing the nucleotide sequences. In a first embodiment, *B. thuringiensis* cells comprising modifications of a nucleotide sequence of this invention are made. Such modifications encompass mutations or deletions of existing regulatory elements, thus leading to altered expression of the nucleotide sequence, or the incorporation of new regulatory elements controlling the expression of the nucleotide sequence. In another embodiment, additional copies of one or more of the nucleotide sequences are added to *Bacillus thuringiensis* cells either by insertion into the chromosome or by introduction of extrachromosomally replicating molecules containing the nucleotide sequences.

[00117] In another embodiment, at least one of the nucleotide sequences of the invention is inserted into an appropriate expression cassette, comprising a promoter and termination signal. Expression of the nucleotide sequence is constitutive, or an inducible promoter responding to various types of stimuli to initiate transcription is used. In a preferred embodiment, the cell in which the toxin is expressed is a microorganism, such as a virus, bacteria, or a fungus. In a preferred embodiment, a virus, such as a baculovirus, contains a nucleotide sequence of the invention in its genome and expresses large amounts of the corresponding insecticidal toxin after infection of appropriate eukaryotic cells that are suitable for virus replication and expression of the nucleotide sequence. The insecticidal toxin thus produced is used as an insecticidal agent. Alternatively, baculoviruses engineered to include the nucleotide sequence are used to infect insects *in vivo* and kill them either by expression of the insecticidal toxin or by a combination of viral infection and expression of the insecticidal toxin.

[00118] Bacterial cells are also hosts for the expression of the nucleotide sequences of the invention. In a preferred embodiment, non-pathogenic symbiotic bacteria, which are able to live and replicate within plant tissues, so-called endophytes, or non- pathogenic symbiotic bacteria, which are capable of colonizing the phyllosphere or the rhizosphere, so-called epiphytes, are used. Such bacteria include bacteria of the genera *Agrobacterium*,

Alcaligenes, Azospirillum, Azotobacter, Bacillus, Clavibacter, Enterobacter, Erwinia, Flavobacter, Klebsiella, Pseudomonas, Rhizobium, Serratia, Streptomyces and *Xanthomonas*. Symbiotic fungi, such as *Trichoderma* and *Gliocladium* are also possible hosts for expression of the inventive nucleotide sequences for the same purpose.

[00119] Techniques for these genetic manipulations are specific for the different available hosts and are known in the art. For example, the expression vectors pKK223-3 and pKK223-2 can be used to express heterologous genes in *E. coli*, either in transcriptional or translational fusion, behind the tac or trc promoter. For the expression of operons encoding multiple ORFs, the simplest procedure is to insert the operon into a vector such as pKK223- 3 in transcriptional fusion, allowing the cognate ribosome binding site of the heterologous genes to be used. Techniques for overexpression in gram-positive species such as *Bacillus* are also known in the art and can be used in the context of this invention (Quax et al. In:Industrial Microorganisms:Basic and Applied Molecular Genetics, Eds. Baltz *et al.*, American Society for Microbiology, Washington (1993)). Alternate systems for overexpression rely for example, on yeast vectors and include the use of Pichia, Saccharomyces and Kluyveromyces (Sreekrishna, In:Industrial microorganisms:basic and applied molecular genetics, Baltz, Hegeman, and Skatrud eds., American Society for Microbiology, Washington (1993); Dequin & Barre, Biotechnology L2:173- 177 (1994); van den Berg *et al.*, Biotechnology 8:135-139 (1990)).

Plant transformation

[00120] In a particularly preferred embodiment, at least one of the insecticidal modified Cry3A toxins of the invention is expressed in a higher organism, e.g., a plant. In this case, transgenic plants expressing effective amounts of the modified Cry3A toxins protect themselves from insect pests. When the insect starts feeding on such a transgenic plant, it also ingests the expressed modified Cry3A toxins. This will deter the insect from further biting into the plant tissue or may even harm or kill the insect. A nucleotide sequence of the present invention is inserted into an expression cassette, which is then preferably stably integrated in the genome of said plant. In another preferred embodiment, the nucleotide sequence is included in a non- pathogenic self- replicating virus. Plants transformed in accordance with the present invention may be monocots or dicots and include, but are not limited to, maize, wheat, barley, rye, sweet potato, bean, pea, chicory,

lettuce, cabbage, cauliflower, broccoli, turnip, radish, spinach, asparagus, onion, garlic, pepper, celery, squash, pumpkin, hemp, zucchini, apple, pear, quince, melon, plum, cherry, peach, nectarine, apricot, strawberry, grape, raspberry, blackberry, pineapple, avocado, papaya, mango, banana, soybean, tomato, sorghum, sugarcane, sugar beet, sunflower, rapeseed, clover, tobacco, carrot, cotton, alfalfa, rice, potato, eggplant, cucumber, *Arabidopsis*, and woody plants such as coniferous and deciduous trees.

[00121] Once a desired nucleotide sequence has been transformed into a particular plant species, it may be propagated in that species or moved into other varieties of the same species, particularly including commercial varieties, using traditional breeding techniques.

[00122] A nucleotide sequence of this invention is preferably expressed in transgenic plants, thus causing the biosynthesis of the corresponding modified Cry3A toxin in the transgenic plants. In this way, transgenic plants with enhanced resistance to insects are generated. For their expression in transgenic plants, the nucleotide sequences of the invention may require other modifications and optimization. Although in many cases genes from microbial organisms can be expressed in plants at high levels without modification, low expression in transgenic plants may result from microbial nucleotide sequences having codons that are not preferred in plants. It is known in the art that all organisms have specific preferences for codon usage, and the codons of the nucleotide sequences described in this invention can be changed to conform with plant preferences, while maintaining the amino acids encoded thereby. Furthermore, high expression in plants is best achieved from coding sequences that have at least about 35% GC content, preferably more than about 45%, more preferably more than about 50%, and most preferably more than about 60%. Microbial nucleotide sequences that have low GC contents may express poorly in plants due to the existence of ATTTA motifs that may destabilize messages, and AATAAA motifs that may cause inappropriate polyadenylation. Although preferred gene sequences may be adequately expressed in both monocotyledonous and dicotyledonous plant species, sequences can be modified to account for the specific codon preferences and GC content preferences of monocotyledons or dicotyledons as these preferences have been shown to differ (Murray *et al.* Nucl. Acids Res. 17:477-498 (1989)). In addition, the nucleotide sequences are screened for the existence of illegitimate splice sites that may cause message truncation. All changes required to be made within the nucleotide sequences such as those described

320
321

above are made using well known techniques of site directed mutagenesis, PCR, and synthetic gene construction using the methods described in the published patent applications EP 0 385 962 (to Monsanto), EP 0 359 472 (to Lubrizol, and WO 93/07278 (to Ciba-Geigy).

[00123] In one embodiment of the invention a *cry*3A gene is made according to the procedure disclosed in U.S. Patent 5,625,136, herein incorporated by reference. In this procedure, maize preferred codons, i.e., the single codon that most frequently encodes that amino acid in maize, are used. The maize preferred codon for a particular amino acid might be derived, for example, from known gene sequences from maize. Maize codon usage for 28 genes from maize plants is found in Murray et al., Nucleic Acids Research 17:477-498 (1989), the disclosure of which is incorporated herein by reference. A synthetic sequence made with maize optimized codons is set forth in SEQ ID NO: 3.

[00124] In this manner, the nucleotide sequences can be optimized for expression in any plant. It is recognized that all or any part of the gene sequence may be optimized or synthetic. That is, synthetic or partially optimized sequences may also be used.

[00125] For efficient initiation of translation, sequences adjacent to the initiating methionine may require modification. For example, they can be modified by the inclusion of sequences known to be effective in plants. Joshi has suggested an appropriate consensus for plants (NAR 15:6643-6653 (1987)) and Clonetech suggests a further consensus translation initiator (1993/1994 catalog, page 210). These consensuses are suitable for use with the nucleotide sequences of this invention. The sequences are incorporated into constructions comprising the nucleotide sequences, up to and including the ATG (whilst leaving the second amino acid unmodified), or alternatively up to and including the GTC subsequent to the ATG (with the possibility of modifying the second amino acid of the transgene).

[00126] Expression of the nucleotide sequences in transgenic plants is driven by promoters that function in plants. The choice of promoter will vary depending on the temporal and spatial requirements for expression, and also depending on the target species. Thus, expression of the nucleotide sequences of this invention in leaves, in stalks or stems, in ears, in inflorescences (e.g. spikes, panicles, cobs, etc.), in roots, and/or seedlings is preferred. In many cases, however, protection against more than one type of insect pest is sought, and thus expression in multiple tissues is desirable. Although many promoters from dicotyledons have been shown to be operational in monocotyledons and vice versa,

ideally dicotyledonous promoters are selected for expression in dicotyledons, and monocotyledonous promoters for expression in monocotyledons. However, there is no restriction to the provenance of selected promoters; it is sufficient that they are operational in driving the expression of the nucleotide sequences in the desired cell.

[00127] Preferred promoters that are expressed constitutively include promoters from genes encoding actin or ubiquitin and the CaMV 35S and 19S promoters. The nucleotide sequences of this invention can also be expressed under the regulation of promoters that are chemically regulated. This enables the insecticidal modified Cry3A toxins to be synthesized only when the crop plants are treated with the inducing chemicals. Preferred technology for chemical induction of gene expression is detailed in the published application EP 0 332 104 (to Ciba- Geigy) and U.S. Patent 5,614,395. A preferred promoter for chemical induction is the tobacco PR-1a promoter.

[00128] A preferred category of promoters is that which is wound inducible. Numerous promoters have been described which are expressed at wound sites and also at the sites of phytopathogen infection. Ideally, such a promoter should only be active locally at the sites of infection, and in this way the insecticidal modified Cry3A toxins only accumulate in cells that need to synthesize the insecticidal modified Cry3A toxins to kill the invading insect pest. Preferred promoters of this kind include those described by Stanford *et al.* Mol. Gen. Genet. 215:200-208 (1989), Xu *et al.* Plant Molec. Biol. 22:573-588 (1993), Logemann *et al.* Plant Cell 1:151-158 (1989), Rohrmeier & Lehle, Plant Molec. Biol. 22:783-792 (1993), Firek *et al.* Plant Molec. Biol. 22:129-142 (1993), and Warner *et al.* *Plant J.* 3:191-201 (1993).

[00129] Tissue-specific or tissue-preferential promoters useful for the expression of the modified Cry3A toxin genes in plants, particularly maize, are those which direct expression in root, pith, leaf or pollen, particularly root. Such promoters, e.g. those isolated from PEPC or trpA, are disclosed in U.S. Pat. No. 5,625,136, or MTL, disclosed in U.S. Pat. No. 5,466,785. Both U. S. patents are herein incorporated by reference in their entirety.

[00130] Further preferred embodiments are transgenic plants expressing the nucleotide sequences in a wound-inducible or pathogen infection-inducible manner.

[00131] In addition to promoters, a variety of transcriptional terminators are also available for use in chimeric gene construction using the modified Cry3A toxin genes of the present invention. Transcriptional terminators are responsible for the termination of transcription

beyond the transgene and its correct polyadenylation. Appropriate transcriptional terminators and those that are known to function in plants include the CaMV 35S terminator, the tml terminator, the nopaline synthase terminator, the pea rbcS E9 terminator and others known in the art. These can be used in both monocotyledons and dicotyledons. Any available terminator known to function in plants can be used in the context of this invention.

[00132] Numerous other sequences can be incorporated into expression cassettes described in this invention. These include sequences that have been shown to enhance expression such as intron sequences (e.g. from Adh1 and bronze1) and viral leader sequences (e.g. from TMV, MCMV and AMV).

[00133] It may be preferable to target expression of the nucleotide sequences of the present invention to different cellular localizations in the plant. In some cases, localization in the cytosol may be desirable, whereas in other cases, localization in some subcellular organelle may be preferred. Subcellular localization of transgene-encoded enzymes is undertaken using techniques well known in the art. Typically, the DNA encoding the target peptide from a known organelle-targeted gene product is manipulated and fused upstream of the nucleotide sequence. Many such target sequences are known for the chloroplast and their functioning in heterologous constructions has been shown. The expression of the nucleotide sequences of the present invention is also targeted to the endoplasmic reticulum or to the vacuoles of the host cells. Techniques to achieve this are well known in the art.

[00134] Vectors suitable for plant transformation are described elsewhere in this specification. For *Agrobacterium*-mediated transformation, binary vectors or vectors carrying at least one T-DNA border sequence are suitable, whereas for direct gene transfer any vector is suitable and linear DNA containing only the construction of interest may be preferred. In the case of direct gene transfer, transformation with a single DNA species or co-transformation can be used (Schocher *et al.* Biotechnology 4:1093- 1096 (1986)). For both direct gene transfer and Agrobacterium-mediated transfer, transformation is usually (but not necessarily) undertaken with a selectable marker that may provide resistance to an antibiotic (kanamycin, hygromycin or methotrexate) or a herbicide (basta). Plant transformation vectors comprising the modified Cry3A toxin genes of the present invention may also comprise genes (e.g. phosphomannose isomerase; PMI) which provide for positive selection of the transgenic plants as disclosed in U.S.

Patents 5,767,378 and 5,994,629, herein incorporated by reference. The choice of selectable marker is not, however, critical to the invention.

[00135] In another embodiment, a nucleotide sequence of the present invention is directly transformed into the plastid genome. A major advantage of plastid transformation is that plastids are generally capable of expressing bacterial genes without substantial codon optimization, and plastids are capable of expressing multiple open reading frames under control of a single promoter. Plastid transformation technology is extensively described in U.S. Patent Nos. 5,451,513, 5,545,817, and 5,545,818, in PCT application no. WO 95/16783, and in McBride *et al.* (1994) Proc. Nati. Acad. Sci. USA 91, 7301-7305. The basic technique for chloroplast transformation involves introducing regions of cloned plastid DNA flanking a selectable marker together with the gene of interest 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. Nati. Acad. Sci. USA 87, 8526-8530; Staub, J. M., and Maliga, P. (1992) Plant Cell 4, 39-45). This resulted in stable homoplasmic transformants at a frequency of approximately one per 100 bombardments of target leaves. 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'- adenyltransf erase (Svab, Z., and Maliga, P. (1993) Proc. Natl. Acad. Sci. USA 90, 913-917). Previously, this marker had been used successfully for high-frequency transformation of the plastid genome of the green alga Chlamydomonas reinhardtii (Goldschmidt- Clermont, M. (1991) Nucl. Acids Res. 19:4083-4089). Other selectable markers useful for plastid transformation are known in the art and encompassed within the scope of the invention. Typically, approximately 15-20 cell division cycles following transformation are required to reach a homoplastidic state. Plastid expression, in which genes are inserted by homologous recombination into

all of 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, a nucleotide sequence of the present invention is inserted into a plastid-targeting vector and transformed into the plastid genome of a desired plant host. Plants homoplastic for plastid genomes containing a nucleotide sequence of the present invention are obtained, and are preferentially capable of high expression of the nucleotide sequence.

Combinations of Insect Control Principles

[00136] The modified Cry3A toxins of the invention can be used in combination with Bt δ-endotoxins or other pesticidal principles to increase pest target range. Furthermore, the use of the modified Cry3A toxins of the invention in combination with Bt δ-endotoxins or other pesticidal principles of a distinct nature has particular utility for the prevention and/or management of insect resistance.

[00137] Other insecticidal principles include, for example, lectins, α-amylase, peroxidase and cholesterol oxidase. Vegetative Insecticidal Protein genes, such as *vip*1A(a) and *vip*2A(a) as disclosed in U.S. Pat. No. 5,889,174 and herein incorporated by reference, are also useful in the present invention.

[00138] This co-expression of more than one insecticidal principle in the same transgenic plant can be achieved by genetically engineering a plant to contain and express all the genes necessary. Alternatively, a plant, Parent 1, can be genetically engineered for the expression of genes of the present invention. A second plant, Parent 2, can be genetically engineered for the expression of a supplemental insect control principle. By crossing Parent 1 with Parent 2, progeny plants are obtained which express all the genes introduced into Parents 1 and 2.

[00139] Transgenic seed of the present invention can also be treated with an insecticidal seed coating as described in U. S. Patent Nos. 5,849,320 and 5,876,739, herein incorporated by reference. Where both the insecticidal seed coating and the transgenic seed of the invention are active against the same target insect, the combination is useful (i) in a method for enhancing activity of a modified Cry3A toxin of the invention against the target insect and (ii) in a method for preventing development of resistance to a

modified Cry3A toxin of the invention by providing a second mechanism of action against the target insect. Thus, the invention provides a method of enhancing activity against or preventing development of resistance in a target insect, for example corn rootworm, comprising applying an insecticidal seed coating to a transgenic seed comprising one or more modified Cry3A toxins of the invention.

[00140] Even where the insecticidal seed coating is active against a different insect, the insecticidal seed coating is useful to expand the range of insect control, for example by adding an insecticidal seed coating that has activity against lepidopteran insects to the transgenic seed of the invention, which has activity against coleopteran insects, the coated transgenic seed produced controls both lepidopteran and coleopteran insect pests.

EXAMPLES

[00141] The invention will be further described by reference to the following detailed examples. These examples are provided for the purposes of illustration only, and are not intended to be limiting unless otherwise specified. Standard recombinant DNA and molecular cloning techniques used here are well known in the art and are described by J. Sambrook, *et al.*, Molecular Cloning: *A Laboratory Manual, 3d Ed.*, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press (2001); by T.J. Silhavy, M.L. Berman, and L.W. Enquist, *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) and by Ausubel, F.M. *et al.*, *Current Protocols in Molecular Biology*, New York, John Wiley and Sons Inc., (1988), Reiter, *et al.*, *Methods in Arabidopsis Research*, World Scientific Press (1992), and Schnitz *et al.*, *Plant Molecular Biology Manual*, Kluwer Academic Publishers (1998).

Example 1. Maize Optimized *cry*3A Gene Construction

[00142] The maize optimized *cry*3A gene is made according to the procedure disclosed in U.S. Patent 5,625,136. In this procedure, maize preferred codons, i.e., the single codon that most frequently encodes that amino acid in maize, are used. The maize preferred codon for a particular amino acid is derived from known gene sequences from maize. Maize codon usage for 28 genes from maize plants is found in Murray *et al.*, Nucleic

Acids Research 17:477-498 (1989). The synthetic sequence made with maize optimized codons is set forth in SEQ ID NO: 3.

Example 2. Identification of Cathepsin-G Enzymatic Activity in Western Corn Rootworm Guts

[00143] Cathepsin G-like (serine protease) and cathepsin B-like (cysteine protease) enzymatic activities in western corn rootworm guts are measured using colorimetric substrates. Each 1 ml reaction contains five homogenized midguts of the 3rd instar of western corn rootworm and 1 mg of substrate dissolved in reaction buffer (10 mM Tris, 5 mM NaCl, 0.01 M DTT, pH 7.5). The cathepsin G substrate tested is Ala-Ala-Pro-Phe (SEQ ID NO: 35)-*p*NA and cathepsin B substrate, Arg-Arg-*p*NA. The reactions are incubated at 28°C for 1hr. The intensity of yellow color formation, indicative of the efficiency of a protease to recognize the appropriate substrate, is compared in treatments vs. controls. The reactions are scored as negative (-) if no color or slight background color is detected. Reactions which are 25%, 50%, 75% or 100% above background are scored as +, ++, +++, or ++++, respectively.

[00144] Results of the enzymatic assays are shown in the following table.

Reaction	Product Color intensity
WCR gut only	-
Cathepsin B substrate only	-
Cathepsin G substrate only	-
WCR gut + Cathepsin B substrate	+
WCR gut + Cathepsin G substrate	+++

[00145] This is the first time that the serine protease cathepsin G activity has been identified in western corn rootworm guts. Western corn rootworm guts clearly have stronger cathepsin G, the serine protease, activity compared to cathepsin B, the cysteine protease, activity. The AAPF sequence (SEQ ID NO: 35) is selected as the cathepsin G protease recognition site for creating modified Cry3A toxins of the present invention.

Example 3. Construction of Modified *cry3A* Genes.

[00146] Modified *cry*3A genes comprising a nucleotide sequence that encodes the cathepsin G recognition site in domain I, domain III, or domain I and domain III are made using overlap PCR. The maize optimized *cry*3A gene (SEQ ID NO: 2), comprised in plasmid pCIB6850 (SEQ ID NO: 5), is used as the starting template. Eight modified *cry*3A gene constructs, which encode modified Cry3A toxins, are made; *cry*3A054, *cry*3A055, and *cry*3A085, which comprise the cathepsin G recognition site coding sequence in domain I; *cry*3A058, *cry*3A082, which comprise the cathepsin G recognition site coding sequence in domain III; *cry*3A056, *cry*3A057, *cry*3A083, which comprise the cathepsin G recognition site coding sequence in domain I and domain III. The eight modified *cry*3A genes and the modified Cry3A toxins they encode are described as follows:

*cry*3A054 comprised in pCMS054

[00147] *cry*3A054 (SEQ ID NO: 6) comprises a nucleotide sequence encoding a modified Cry3A toxin. Three overlap PCR primer pairs are used to insert the nucleotide sequence encoding the cathepsin G recognition site into the unmodified maize optimized *cry*3A:

1. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3' (SEQ ID NO: 22)
   AAPFtail3 – 5'-GAACGGTGCAGCGGGGTTCTTCTGCCAGC-3' (SEQ ID NO: 23)

2. Tail5mod – 5'-GCTGCACCGTTCCCCCACAGCCAGGGCCG-3' (SEQ ID NO: 24)
   XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3' (SEQ ID NO: 25)

3. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3' (SEQ ID NO: 22)
   XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3' (SEQ ID NO: 25)

[00148] Primer pair 1 and primer pair 2 generate two unique PCR products. These products are then combined in equal parts and primer pair 3 is used to join the products to generate one PCR fragment that is cloned back into the original pCIB6850 template. The modified *cry*3A054 gene is then transferred to pBluescript (Stratagene). The resulting plasmid is designated pCMS054 and comprises the *cry*3A054 gene (SEQ ID NO: 6).

[00149] The modified Cry3A054 toxin (SEQ ID NO: 7), encoded by the modified *cry*3A gene comprised in pCMS054, has a cathepsin G recognition site, comprising the amino acid sequence AAPF (SEQ ID NO: 35), inserted in domain I between amino acids 107

and 113 of the unmodified Cry3A toxin (SEQ ID NO: 4). The cathepsin G recognition site replaces the naturally occurring trypsin recognition site and is adjacent to a naturally occurring chymotrypsin recognition site.

*cry*3A055 comprised in pCMS055

[00150] *cry*3A055 (SEQ ID NO: 8) comprises a nucleotide sequence encoding a modified Cry3A toxin. Three overlap PCR primer pairs are used to insert the nucleotide sequence encoding the cathepsin G recognition site into the unmodified maize optimized *cry*3A:

1. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3' (SEQ ID NO: 22)
   AAPFtail3 – 5'-GAACGGTGCAGCGGGGTTCTTCTGCCAGC-3' (SEQ ID NO: 23)

2. AAPFtail4 – 5'-GCTGCACCGTTCCGCAACCCCCACAGCCA-3' (SEQ ID NO: 26)
   XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3' (SEQ ID NO: 25)

3. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3' (SEQ ID NO: 22)
   XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3' (SEQ ID NO: 25)

[00151] Primer pair 1 and primer pair 2 generate two unique PCR products. These products are then combined in equal parts and primer pair 3 is used to join the products to generate one PCR fragment that is cloned back into the original pCIB6850 template. The modified *cry*3A055 gene is then transferred to pBluescript (Stratagene). The resulting plasmid is designated pCMS055 and comprises the *cry*3A055 gene (SEQ ID NO: 8).

[00152] The modified Cry3A055 toxin (SEQ ID NO: 9), encoded by the modified *cry*3A gene comprised in pCMS055, has a cathepsin G recognition site, comprising the amino acid sequence AAPF (SEQ ID NO: 35), inserted in domain I between amino acids 107 and 111 of the unmodified Cry3A toxin (SEQ ID NO: 4). The cathepsin G recognition site is adjacent to a natural trypsin and chymotrypsin recognition site.

*cry*3A058 comprised in pCMS058

[00153] *cry*3A058 (SEQ ID NO: 14) comprises a nucleotide sequence encoding a modified Cry3A toxin. Three overlap PCR primer pairs are used to insert the nucleotide

sequence encoding the cathepsin G recognition site into the unmodified maize optimized *cry*3A:

1. SalExt – 5'-GAGCGTCGACTTCTTCAAC-3' (SEQ ID NO: 27)
   AAPF-Y2 – 5'-GAACGGTGCAGCGTATTGGTTGAAGGGGGC-3' (SEQ ID NO: 28)

2. AAPF-Y1 – 5'-GCTGCACCGTTCTACTTCGACAAGACCATC-3' (SEQ ID NO: 29)
   SacExt – 5'-GAGCTCAGATCTAGTTCACGG-3' (SEQ ID NO: 30)

3. SalExt – 5'-GAGCGTCGACTTCTTCAAC-3' (SEQ ID NO: 27)
   SacExt – 5'-GAGCTCAGATCTAGTTCACGG-3' (SEQ ID NO: 30)

[00154] Primer pair 1 and primer pair 2 generate two unique PCR products. These products are then combined in equal parts and primer pair 3 is used to join the products to generate one PCR fragment that is cloned back into the original pCIB6850 template. The modified *cry*3A058 gene is then transferred to pBluescript (Stratagene). The resulting plasmid is designated pCMS058 and comprises the *cry*3A058 gene (SEQ ID NO: 14).

[00155] The modified Cry3A058 toxin (SEQ ID NO: 15), encoded by the modified *cry*3A gene, has a cathepsin G recognition site, comprising the amino acid sequence AAPF (SEQ ID NO: 35), inserted in domain III between amino acids 540and 541 of the unmodified Cry3A toxin (SEQ ID NO: 4). The cathepsin G recognition site is within a naturally occurring chymotrypsin recognition site.

pCMS082 comprising *cry*3A082

[00156] *cry*3A082 (SEQ ID NO: 12) comprises a nucleotide sequence encoding a modified Cry3A toxin. A QuikChange Site Directed Mutagenesis PCR primer pair is used to insert the nucleotide sequence encoding the cathepsin G recognition site into the unmodified maize optimized *cry*3A:

BBmod1 – 5'-CGGGGCCCCCGCTGCACCGTTCTACTTCGACA-3' (SEQ ID NO: 31)
BBmod2 – 5'-TGTCGAAGTAGAACGGTGCAGCGGGGGCCCCG-3'(SEQ ID NO: 32)

[00157] The primer pair generates a unique PCR product. This product is cloned back into the original pCIB6850 template. The modified *cry*3A082 gene is then transferred to

pBluescript (Stratagene). The resulting plasmid is designated pCMS082 and comprises the *cry*3A082 gene (SEQ ID NO: 12).

[00158] The modified Cry3A082 toxin (SEQ ID NO: 13), encoded by the modified *cry*3A gene, has a cathepsin G recognition site, comprising the amino acid sequence AAPF (SEQ ID NO: 35), inserted in domain III between amino acids 539 and 542 of the unmodified Cry3A toxin (SEQ ID NO: 4). The cathepsin G recognition site replaces a naturally occurring chymotrypsin recognition site.

*cry*3A056 comprised in pCMS056

*cry*3A056 (SEQ ID NO: 18) comprises a nucleotide sequence encoding a modified Cry3A toxin. Six overlap PCR primer pairs are used to insert two cathepsin G recognition sites into the unmodified *cry*3A:

1. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3' (SEQ ID NO: 22)
   AAPFtail3 – 5'-GAACGGTGCAGCGGGGTTCTTCTGCCAGC-3' (SEQ ID NO: 23)

2. AAPFtail4 – 5'-GCTGCACCGTTCCGCAACCCCCACAGCCA-3' (SEQ ID NO: 26)
   XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3' (SEQ ID NO: 25)

3. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3' (SEQ ID NO: 22)
   XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3' (SEQ ID NO: 25)

4. SalExt – 5'-GAGCGTCGACTTCTTCAAC-3' (SEQ ID NO: 27)
   AAPF-Y2 – 5'-GAACGGTGCAGCGTATTGGTTGAAGGGGGC-3' (SEQ ID NO: 28)

5. AAPF-Y1 – 5'-GCTGCACCGTTCTACTTCGACAAGACCATC-3' (SEQ ID NO: 29)
   SacExt – 5'-GAGCTCAGATCTAGTTCACGG-3' (SEQ ID NO: 30)

6. SalExt – 5'-GAGCGTCGACTTCTTCAAC-3' (SEQ ID NO: 27)
   SacExt – 5'-GAGCTCAGATCTAGTTCACGG-3' (SEQ ID NO: 30)

[00159] Primer pair 1 and primer pair 2 generate two unique PCR products. These products are combined in equal parts and primer pair 3 is used to join the products to generate one PCR fragment that is cloned back into the original pCIB6850 plasmid. The modified *cry*3A055 gene is then transferred to pBluescript (Stratagene). The resulting plasmid is designated pCMS055. Primer pair 4 and primer pair 5 generate another unique set of fragments that are joined by another PCR with primer pair 6. This fragment is cloned into domain III of the modified *cry*3A055 gene comprised in pCMS055. The

resulting plasmid is designated pCMS056 and comprises the *cry*3A056 gene (SEQ ID NO: 18).

[00160] The modified Cry3A056 toxin (SEQ ID NO: 19), encoded by the modified *cry*3A gene, has a cathepsin G recognition site, comprising the amino acid sequence AAPF (SEQ ID NO: 35), inserted in domain I between amino acids 107 and 111 and in domain III between amino acids 540 and 541 of the unmodified Cry3A toxin (SEQ ID NO: 4). The cathepsin G recognition site is adjacent to a naturally occurring trypsin and chymotrypsin recognition site in domain I and is within a naturally occurring chymotrypain recognition site in domain III.

*cry*3A057 comprised in pCMS057

[00161] *cry*3A057 (SEQ ID NO: 16) comprises a nucleotide sequence encoding a modified Cry3A toxin. Six overlap PCR primer pairs are used to insert two cathepsin G recognition sites into the unmodified *cry*3A:

1. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3' (SEQ ID NO: 22)
   AAPFtail3 – 5'-GAACGGTGCAGCGGGGTTCTTCTGCCAGC-3' (SEQ ID NO: 23)

2. Tail5mod – 5'-GCTGCACCGTTCCCCCACAGCCAGGGCCG-3' (SEQ ID NO: 24)
   XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3' (SEQ ID NO: 25)

3. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3' (SEQ ID NO: 22)
   XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3' (SEQ ID NO: 25)

4. SalExt – 5'-GAGCGTCGACTTCTTCAAC-3' (SEQ ID NO: 27)
   AAPF-Y2 – 5'-GAACGGTGCAGCGTATTGGTTGAAGGGGGC-3' (SEQ ID NO: 28)

5. AAPF-Y1 – 5'-GCTGCACCGTTCTACTTCGACAAGACCATC-3' (SEQ ID NO: 29)
   SacExt – 5'-GAGCTCAGATCTAGTTCACGG-3' (SEQ ID NO: 30)

6. SalExt – 5'-GAGCGTCGACTTCTTCAAC-3' (SEQ ID NO: 27)
   SacExt – 5'-GAGCTCAGATCTAGTTCACGG-3' (SEQ ID NO: 30)

[00162] Primer pair 1 and primer pair 2 generate two unique PCR products. These products are combined in equal parts and primer pair 3 is used to join the products to generate one PCR fragment that is cloned back into the original pCIB6850 plasmid. The modified *cry*3A054 gene is then transferred to pBluescript (Stratagene). The resulting plasmid is designated pCMS054. Primer pair 4 and primer pair 5 generate another unique

set of fragments that are joined by another PCR with primer pair 6. This fragment is cloned into domain III of the modified *cry*3A054 gene comprised in pCMS054. The resulting plasmid is designated pCMS057 and comprises the *cry*3A057 gene (SEQ ID NO: 16).

[00163] The modified Cry3A057 toxin (SEQ ID NO: 17), encoded by the modified *cry*3A gene, has a cathepsin G recognition site, comprising the amino acid sequence AAPF (SEQ ID NO: 35), inserted in domain I between amino acids 107 and 113 and in domain III between amino acids 540 and 541 of the unmodified Cry3A toxin (SEQ ID NO: 4). The cathepsin G recognition site replaces a naturally occurring trypsin recognition site and is adjacent to a naturally occurring chymotrypain recognition site in domain I and is within a naturally occuring chymotrypain recognition site in domain III.

*cry*3A083 comprised in pCMS083

[00164] *cry*3A083 (SEQ ID NO: 20) comprises a nucleotide sequence encoding a modified Cry3A toxin. Three overlap PCR primer pairs and one QuikChange Site Directed Mutagenesis PCR primer pair are used to insert two cathepsin G recognition sites into the unmodified *cry*3A:

1. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3' (SEQ ID NO: 22)
   AAPFtail3 – 5'-GAACGGTGCAGCGGGGTTCTTCTGCCAGC-3' (SEQ ID NO: 23)

2. AAPFtail4 – 5'-GCTGCACCGTTCCGCAACCCCCACAGCCA-3' (SEQ ID NO: 26)
   XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3' (SEQ ID NO: 25)

3. BamExt1 – 5'-GGATCCACCATGACGGCCGAC-3' (SEQ ID NO: 22)
   XbaIExt2 – 5'-TCTAGACCCACGTTGTACCAC-3' (SEQ ID NO: 25)

BBmod1 – 5'-CGGGGCCCCCGCTGCACCGTTCTACTTCGACA-3 (SEQ ID NO: 31)
BBmod2 – 5'-TGTCGAAGTAGAACGGTGCAGCGGGGGCCCCG-3' (SEQ ID NO: 32)

[00165] Primer pair 1 and primer pair 2 generate two unique PCR products. These products are combined in equal parts and primer pair 3 is used to join the products to generate one PCR fragment that is cloned back into the original pCIB6850 plasmid. The modified *cry*3A055 gene is then transferred to pBluescript (Stratagene). The resulting plasmid is designated pCMS055. Primer pair 4 generates another unique fragment that is

cloned into domain III of the modified *cry*3A comprised in pCMS055. The resulting plasmid is designated pCMS083 and comprises the *cry*3A083 gene (SEQ ID NO: 20).

[00166] The modified Cry3A083 toxin (SEQ ID NO: 21), encoded by the modified *cry*3A gene, has a cathepsin G recognition site, comprising the amino acid sequence AAPF (SEQ ID NO: 35), inserted in domain I between amino acids 107 and 111 and between amino acids 539 and 542 of the unmodified Cry3A toxin (SEQ ID NO: 4). The cathepsin G recognition site is adjacent to a naturally occurring trypsin and chymotrypsin recognition site in domain I and replaces a naturally occurring chymotrypsin recognition site in domain III.

*cry*3A085 comprised in pCMS085

[00167] The *cry*3A085 gene (SEQ ID NO: 10) comprises a cathepsin G coding sequence at the same position as in the *cry*3A055 gene described above. The *cry*3A085 gene has an additional 24 nucleotides inserted at the 5' end which encode amino acids 41-47 of the deduced amino acid sequence set forth in SEQ ID NO: 2 as well as an additional methionine. The additional nucleotides are inserted at the 5' end of the *cry*3A055 gene using the following PCR primer pair:

mo3Aext- 5'-GGATCCACCATGAACTACAAGGAGTTCCTCCGCATGACCGC CGACAAC-3' (SEQ ID NO: 33)
CMS16 – 5'-CCTCCACCTGCTCCATGAAG-3' (SEQ ID NO: 34)

[00168] The modified Cry3A085 toxin (SEQ ID NO: 11), encoded by the modified *cry*3A gene, has a cathepsin G recognition site, comprising the amino acid sequence AAPF (SEQ ID NO: 35), inserted in domain I between amino acids corresponding to 107 and 111 of the unmodified Cry3A toxin (SEQ ID NO: 4) and has an additional eight amino acid residues at the N-terminus of which the second residue corresponds to amino acid number 41 of the amino acid sequence set forth in SEQ ID NO: 2.

Example 4. Insecticidal Activity of Modified Cry3A Toxins

[00169] Modified Cry3A toxins are tested for insecticidal activity against western corn rootworm, northern corn rootworm and southern corn rootworm in insect bioassays.

Bioassays are performed using a diet incorporation method. *E. coli* clones that express one of the modified Cry3A toxins of the invention are grown overnight. 500 µl of an overnight culture is sonicated and then mixed with 500 µl of molten artificial diet (Marrone et al. (1985) J. of Economic Entomology 78:290-293). Once the diet solidifies, it is dispensed in a petri-dish and 20 neonate corn rootworm are placed on the diet. The petri-dishes are held at 30°C. Mortality is recorded after 6 days. All of the modified Cry3A toxins cause 50%-100% mortality to western and northern corn rootworm whereas the unmodified Cry3A toxin causes 0%-30% mortality. None of the modified Cry3A toxins have activity against southern corn rootworm.

Example 5. Creation of Transgenic Maize Plants Comprising Modified *cry*3A Coding Sequences

[00170] Three modified cry3A genes, *cry*3A055, representative of a domain I modification, *cry*3A058, representative of a domain III modification, and *cry*3A056, representative of a domain I and domain III modification, are chosen for transformation into maize plants. An expression cassette comprising a modified *cry*3A coding sequence is transferred to a suitable vector for Agrobacterium-mediated maize transformation. For this example, an expression cassette comprises, in addition to the modified *cry*3A gene, the MTL promoter (U.S. Pat. No. 5,466,785) and the nos terminator which is known in the art.

[00171] Transformation of immature maize embryos is performed essentially as described in Negrotto *et al.*, 2000, Plant Cell Reports 19: 798-803. For this example, all media constituents are as described in Negrotto *et al.*, *supra*. However, various media constituents known in the art may be substituted.

[00172] The genes used for transformation are cloned into a vector suitable for maize transformation. Vectors used in this example contain the phosphomannose isomerase (PMI) gene for selection of transgenic lines (Negrotto *et al.* (2000) Plant Cell Reports 19: 798-803).

[00173] *Agrobacterium* strain LBA4404 (pSB1) containing the plant transformation plasmid is grown on YEP (yeast extract (5 g/L), peptone (10g/L), NaCl (5g/L), 15g/l agar, pH 6.8) solid medium for 2 – 4 days at 28°C. Approximately 0.8X 10^9 *Agrobacterium* are

suspended in LS-inf media supplemented with 100 µM As (Negrotto *et al.*,(2000) Plant Cell Rep 19: 798-803). Bacteria are pre-induced in this medium for 30-60 minutes.

[00174] Immature embryos from A188 or other suitable genotype are excised from 8 – 12 day old ears into liquid LS-inf + 100 µM As. Embryos are rinsed once with fresh infection medium. *Agrobacterium* solution is then added and embryos are vortexed for 30 seconds and allowed to settle with the bacteria for 5 minutes. The embryos are then transferred scutellum side up to LSAs medium and cultured in the dark for two to three days. Subsequently, between 20 and 25 embryos per petri plate are transferred to LSDc medium supplemented with cefotaxime (250 mg/l) and silver nitrate (1.6 mg/l) and cultured in the dark for 28°C for 10 days.

[00175] Immature embryos, producing embryogenic callus are transferred to LSD1M0.5S medium. The cultures are selected on this medium for 6 weeks with a subculture step at 3 weeks. Surviving calli are transferred to Reg1 medium supplemented with mannose. Following culturing in the light (16 hour light/ 8 hour dark regiment), green tissues are then transferred to Reg2 medium without growth regulators and incubated for 1-2 weeks. Plantlets are transferred to Magenta GA-7 boxes (Magenta Corp, Chicago Ill.) containing Reg3 medium and grown in the light. After 2-3 weeks, plants are tested for the presence of the PMI genes and the modified cry3A genes by PCR. Positive plants from the PCR assay are transferred to the greenhouse and tested for resistance to corn rootworm.

Example 6. Analysis of Transgenic Maize Plants

Corn Rootworm Efficacy

Root Excision Bioassay

[00176] Plants are sampled as they are being transplanted from Magenta GA-7 boxes into soil. This allows the roots to be sampled from a reasonably sterile environment relative to soil conditions. Sampling consists of cutting a small piece of root (ca. 2-4 cm long) and placing it onto enriched phytagar (phytagar, 12 g., sucrose, 9 g., MS salts, 3 ml., MS vitamins, 3 ml., Nystatin(25mg/ml), 3 ml., Cefotaxime (50mg/ml), 7 ml., Aureomycin (50 mg/ml), 7 ml., Streptomycin (50mg/ml), 7 ml., dH_2O, 600 ml) in a small petri-dish. Negative controls are either transgenic plants that are PCR negative for the modified *cry*3A gene from the same experiment, or from non-transgenic plants (of a similar size to

test plants) that are being grown in the phytotron. If sampling control roots from soil, the root samples are washed with water to remove soil residue, dipped in Nystatin solution (5mg/ml), removed from the dip, blotted dry with paper toweling, and placed into a phytagar dish.

[00177] Root samples are inoculated with western corn rootworms by placing 10 first instar larvae onto the inside surface of the lid of each phytagar dish and the lids then tightly resealed. Larvae are handled using a fine tip paintbrush. After all dishes are inoculated, the tray of dishes is placed in the dark at room temperature until data collection.

[00178] At 3-4 days post inoculation, data is collected. The percent mortality of the larvae is calculated along with a visual damage rating of the root. Feeding damage is rated as high, moderate, low, or absent and given a numerical value of 3, 2, 1 or 0, respectively. Root samples causing at least 40% mortality and having a damage rating of 2 or less are considered positive.

[00179] Results in the following table show that plants expressing a modified Cry3A toxin cause from 40-100% mortality to western corn rootworm whereas control plants cause 0-30% mortality. Also, plants expressing a modified Cry3A toxin sustain significantly less feeding damage than control plants.

T0 Event	Modified Cry3A Toxin Expressed	Percent Mortality Per Plant					Mean Damage Rating Per Event
		A	B	C	D	E	
240A7	Cry3A055	80	40	80	60		0.8
240B2	Cry3A055	60	60	60	80		1.25
240B9	Cry3A055	40	60	60	100		1
240B10	Cry3A055	80	40	60	60		1
240A15	Cry3A055	80	60	50	70	70	0.6
240A5	Cry3A055	60	80	60			0.33
240A9	Cry3A055	50	60	60	70	70	1.6
244A4	Cry3A058	50					1
244A7	Cry3A058	40	40	60			1.3
244A5	Cry3A058	50					1
244B7	Cry3A058	90					1
244B6	Cry3A058	50	40	60			1
243A3	Cry3A056	50	90	80	60		1.25
243A4	Cry3A056	50	80	60			1.7
243B1	Cry3A056	80	90				0.5
243B4	Cry3A056	70	60	50	80		1.5

245B2	Cry3A056	90 50 70 60	1
WT1	-	0 10 20 10 0	2.6
WT2	-	0 30 0 0 20	2.8

Whole Plant Bioassay

[00180] Some positive plants identified using the root excision bioassay described above are evaluated for western corn rootworm resistance using a whole plant bioassay. Plants are infested generally within 3 days after the root excision assay is completed.

[00181] Western corn rootworm eggs are preincubated so that hatch occurs 2-3 days after plant inoculation. Eggs are suspended in 0.2% agar and applied to the soil around test plants at approximately 200 eggs/plant.

[00182] Two weeks after the eggs hatch, plants are evaluated for damage caused by western corn rootworm larvae. Plant height attained, lodging, and root mass are criteria used to determine if plants are resistant to western corn rootworm feeding damage. At the time of evaluation, control plants typically are smaller than modified Cry3A plants. Also, non-transgenic control plants and plants expressing the unmodified Cry3A toxin encoded by the maize optimized *cry*3A gene have lodged during this time due to severe pruning of most of the roots resulting in no root mass accumulation. At the time of evaluation, plants expressing a modified Cry3A toxin of the invention are taller than control plants, have not lodged, and have a large intact root mass due to the insecticidal activity of the modified Cry3A toxin.

ELISA Assay

[00183] ELISA analysis according to the method disclosed in U.S. Patent No. 5,625,136 is used for the quantitative determination of the level of modified and unmodified Cry3A protein in transgenic plants.

Table 3. Whole Plant Bioassay Results and Protein Levels

Transgenic Maize Plant	Type of Cry3A Toxin Expressed	Cry3A Protein Level in Roots (ng/mg)	Plant Lodged	Intact Root Mass
240A2E	modified Cry3A055	224	-	+
240A9C	modified Cry3A055	71	-	+
240B9D	modified Cry3A055	204	-	+
240B9E	modified Cry3A055	186	-	+
240B10D	modified Cry3A055	104	-	+
240B10E	modified Cry3A055	70	-	+
240A15E	modified Cry3A055	122	-	+
240B4D	modified Cry3A055	97	-	+
243B5A	modified Cry3A056	41	-	+
244A7A	modified Cry3A058	191	-	+
710-2-51	maize optimized	39	+	-
710-2-54	maize optimized	857	+	-
710-2-61	maize optimized	241	+	-
710-2-67	maize optimized	1169	+	-
710-2-68	maize optimized	531	+	-
710-2-79	maize optimized	497	+	-
710-2-79	maize optimized	268	+	-
WT1 Control	-	0	+	-
WT2 Control	-	0	+	-

ABSTRACT

Compositions and methods for controlling plant pests are disclosed. In particular, novel nucleic acid sequences encoding modified Cry3A toxins having increased toxicity to corn rootworm are provided. By inserting a protease recognition site, such as cathepsin G, that is recognized by a gut protease of a target insect in at least one position of a Cry3A toxin a modified Cry3A toxin having significantly greater toxicity, particularly to western and northern corn rootworm is designed. Further, a method of making the modified Cry3A toxins and methods of using the modified *cry*3A nucleic acid sequences, for example in microorganisms to control insects or in transgenic plants to confer protection from insect damage, and a method of using the modified Cry3A toxins, and compositions and formulations comprising the modified Cry3A toxins, for example applying the modified Cry3A toxins or compositions or formulations to insect-infested areas, or to prophylactically treat insect-susceptible areas or plants to confer protection against the insect pests are disclosed.

Modified Cry3A Toxins.ST25
SEQUENCE LISTING

<110> Chen, Eric
Stacy, Cheryl

<120> Modified Cry3A Toxins and Nucleic Acid Sequences Coding Therefor

<130> 60065

<160> 38

<170> PatentIn version 3.0

<210> 1
<211> 1932
<212> DNA
<213> Bacillus thuringiensis

<220>
<221> CDS
<222> (1)..(1932)
<223> Native cry3A coding sequence according to Sekar et al. 1987, Proc. Natl. Aca. Sci. 84:7036-7040

<400> 1

```
atg aat ccg aac aat cga agt gaa cat gat aca ata aaa act act gaa       48
Met Asn Pro Asn Asn Arg Ser Glu His Asp Thr Ile Lys Thr Thr Glu
1               5                  10                  15

aat aat gag gtg cca act aac cat gtt caa tat cct tta gcg gaa act       96
Asn Asn Glu Val Pro Thr Asn His Val Gln Tyr Pro Leu Ala Glu Thr
            20                  25                  30

cca aat cca aca cta gaa gat tta aat tat aaa gag ttt tta aga atg      144
Pro Asn Pro Thr Leu Glu Asp Leu Asn Tyr Lys Glu Phe Leu Arg Met
        35                  40                  45

act gca gat aat aat acg gaa gca cta gat agc tct aca aca aaa gat      192
```

Modified Cry3A Toxins.ST25

```
Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys Asp
 50                  55                  60

gtc att caa aaa ggc att tcc gta gta ggt gat ctc cta ggc gta gta      240
Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val Val
65                  70                  75                  80

ggt ttc ccg ttt ggt gga gcg ctt gtt tcg ttt tat aca aac ttt tta      288
Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe Leu
                85                  90                  95

aat act att tgg cca agt gaa gac ccg tgg aag gct ttt atg gaa caa      336
Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu Gln
            100                 105                 110

gta gaa gca ttg atg gat cag aaa ata gct gat tat gca aaa aat aaa      384
Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn Lys
        115                 120                 125

gct ctt gca gag tta cag ggc ctt caa aat aat gtc gaa gat tat gtg      432
Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr Val
    130                 135                 140

agt gca ttg agt tca tgg caa aaa aat cct gtg agt tca cga aat cca      480
Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn Pro
145                 150                 155                 160

cat agc cag ggg cgg ata aga gag ctg ttt tct caa gca gaa agt cat      528
His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser His
                165                 170                 175

ttt cgt aat tca atg cct tcg ttt gca att tct gga tac gag gtt cta      576
Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val Leu
            180                 185                 190

ttt cta aca aca tat gca caa gct gcc aac aca cat tta ttt tta cta      624
Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu Leu
        195                 200                 205

aaa gac gct caa att tat gga gaa gaa tgg gga tac gaa aaa gaa gat      672
Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu Asp
    210                 215                 220

att gct gaa ttt tat aaa aga caa cta aaa ctt acg caa gaa tat act      720
Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr Thr
225                 230                 235                 240

gac cat tgt gtc aaa tgg tat aat gtt gga tta gat aaa tta aga ggt      768
Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg Gly
                245                 250                 255

tca tct tat gaa tct tgg gta aac ttt aac cgt tat cgc aga gag atg      816
Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu Met
            260                 265                 270

aca tta aca gta tta gat tta att gca cta ttt cca ttg tat gat gtt      864
Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp Val
        275                 280                 285

cgg cta tac cca aaa gaa gtt aaa acc gaa tta aca aga gac gtt tta      912
Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val Leu
    290                 295                 300
```

```
aca gat cca att gtc gga gtc aac aac ctt agg ggc tat gga aca acc      960
Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr Thr
305                 310                 315                 320

ttc tct aat ata gaa aat tat att cga aaa cca cat cta ttt gac tat     1008
Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp Tyr
                325                 330                 335

ctg cat aga att caa ttt cac acg cgg ttc caa cca gga tat tat gga     1056
Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr Gly
            340                 345                 350

aat gac tct ttc aat tat tgg tcc ggt aat tat gtt tca act aga cca     1104
Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg Pro
        355                 360                 365

agc ata gga tca aat gat ata atc aca tct cca ttc tat gga aat aaa     1152
Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn Lys
    370                 375                 380

tcc agt gaa cct gta caa aat tta gaa ttt aat gga gaa aaa gtc tat     1200
Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val Tyr
385                 390                 395                 400

aga gcc gta gca aat aca aat ctt gcg gtc tgg ccg tcc gct gta tat     1248
Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val Tyr
                405                 410                 415

tca ggt gtt aca aaa gtg gaa ttt agc caa tat aat gat caa aca gat     1296
Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr Asp
            420                 425                 430

gaa gca agt aca caa acg tac gac tca aaa aga aat gtt ggc gcg gtc     1344
Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala Val
        435                 440                 445

tgc tgg gat tct atc gat caa ttg cct cca gaa aca aca gat gaa cct     1392
Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu Pro
    450                 455                 460

tta gaa aag gga tat agc cat caa ctc aat tat gta atg tgc ttt tta     1440
Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe Leu
465                 470                 475                 480

atg cag ggt agt aga gga aca atc cca gtg tta act tgg aca cat aaa     1488
Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His Lys
                485                 490                 495

agt gta gac ttt ttt aac atg att gat tcg aaa aaa att aca caa ctt     1536
Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln Leu
            500                 505                 510

tcg tta gta aag gca tat aag tta caa tct ggt gct tcc gtt gtc gca     1584
Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val Ala
        515                 520                 525

ggt cct agg ttt aca gga gga gat atc att caa tgc aca gaa aat gga     1632
Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn Gly
    530                 535                 540

agt gcg gca act att tac gtt aca ccg gat gtg tcg tac tct caa aaa     1680
Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln Lys
545                 550                 555                 560
```

```
tat cga gct aga att cat tat gct tct aca tct cag ata aca ttt aca      1728
Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe Thr
                565                 570                 575

ctc agt tta gac ggg gca cca ttt aat caa tac tat ttc gat aaa acg      1776
Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp Lys Thr
            580                 585                 590

ata aat aaa gga gac aca tta acg tat aat tca ttt aat tta gca agt      1824
Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala Ser
        595                 600                 605

ttc agc aca cca ttc gaa tta tca ggg aat aac tta caa ata ggc gtc      1872
Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly Val
    610                 615                 620

aca gga tta agt gct gga gat aaa gtt tat ata gac aaa att gaa ttt      1920
Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu Phe
625                 630                 635                 640

att cca gtg aat                                                     1932
Ile Pro Val Asn
```

<210> 2

<211> 644

<212> PRT

<213> Bacillus thuringiensis

<400> 2

```
Met Asn Pro Asn Asn Arg Ser Glu His Asp Thr Ile Lys Thr Thr Glu
 1               5                  10                  15

Asn Asn Glu Val Pro Thr Asn His Val Gln Tyr Pro Leu Ala Glu Thr
            20                  25                  30

Pro Asn Pro Thr Leu Glu Asp Leu Asn Tyr Lys Glu Phe Leu Arg Met
        35                  40                  45

Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys Asp
    50                  55                  60

Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val Val
65                  70                  75                  80

Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe Leu
                85                  90                  95

Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu Gln
```


10/30/2012 Document1

100 105 110

Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn Lys
115 120 125

Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr Val
130 135 140

Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn Pro
145 150 155 160

His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser His
165 170 175

Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val Leu
180 185 190

Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu Leu
195 200 205

Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu Asp
210 215 220

Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr Thr
225 230 235 240

Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg Gly
245 250 255

Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu Met
260 265 270

Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp Val
275 280 285

Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val Leu
290 295 300

Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr Thr
305 310 315 320

Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp Tyr
325 330 335

Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr Gly
340 345 350

EDITED COPY. DO NOT ENTER

Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg Pro
355 360 365

Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn Lys
370 375 380

Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val Tyr
385 390 395 400

Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val Tyr
405 410 415

Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr Asp
420 425 430

Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala Val
435 440 445

Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu Pro
450 455 460

Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe Leu
465 470 475 480

Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His Lys
485 490 495

Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln Leu
500 505 510

Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val Ala
515 520 525

Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn Gly
530 535 540

Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln Lys
545 550 555 560

Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe Thr
565 570 575

Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp Lys Thr
580 585 590

Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala Ser
595 600 605

03TB-02KA.USAT01

Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly Val
610 615 620

Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu Phe
625 630 635 640

Ile Pro Val Asn

<210> 3

<211> 1803

<212> DNA

<213> Artificial Sequence

<220>

<221> CDS

<222> (1)..(1794)

<223> Maize optimized cry3A coding sequence

<400> 3

```
atg acg gcc gac aac aac acc gag gcc ctg gac agc agc acc acc aag      48
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
 1               5                  10                  15

gac gtg atc cag aag ggc atc agc gtg gtg ggc gac ctg ctg ggc gtg      96
Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
            20                  25                  30

gtg ggc ttc ccc ttc ggc ggc gcc ctg gtg agc ttc tac acc aac ttc     144
Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
        35                  40                  45

ctg aac acc atc tgg ccc agc gag gac ccc tgg aag gcc ttc atg gag     192
Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
    50                  55                  60

cag gtg gag gcc ctg atg gac cag aag atc gcc gac tac gcc aag aac     240
Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65                  70                  75                  80

aag gca ctg gcc gag cta cag ggc ctc cag aac aac gtg gag gac tat     288
Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
                85                  90                  95

gtg agc gcc ctg agc agc tgg cag aag aac ccc gtc tcg agc cgc aac     336
Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn
            100                 105                 110

ccc cac agc cag ggc cgc atc cgc gag ctg ttc agc cag gcc gag agc     384
```

Modified Cry3A Toxins.ST25

```
Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser
       115                 120                 125

cac ttc cgc aac agc atg ccc agc ttc gcc atc agc ggc tac gag gtg      432
His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val
   130                 135                 140

ctg ttc ctg acc acc tac gcc cag gcc gcc aac acc cac ctg ttc ctg      480
Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu
145                 150                 155                 160

ctg aag gac gcc caa atc tac gga gag gag tgg ggc tac gag aag gag      528
Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu
                165                 170                 175

gac atc gcc gag ttc tac aag cgc cag ctg aag ctg acc cag gag tac      576
Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr
            180                 185                 190

acc gac cac tgc gtg aag tgg tac aac gtg ggt cta gac aag ctc cgc      624
Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg
        195                 200                 205

ggc agc agc tac gag agc tgg gtg aac ttc aac cgc tac cgc cgc gag      672
Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu
    210                 215                 220

atg acc ctg acc gtg ctg gac ctg atc gcc ctg ttc ccc ctg tac gac      720
Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp
225                 230                 235                 240

gtg cgc ctg tac ccc aag gag gtg aag acc gag ctg acc cgc gac gtg      768
Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val
                245                 250                 255

ctg acc gac ccc atc gtg ggc gtg aac aac ctg cgc ggc tac ggc acc      816
Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr
            260                 265                 270

acc ttc agc aac atc gag aac tac atc cgc aag ccc cac ctg ttc gac      864
Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp
        275                 280                 285

tac ctg cac cgc atc cag ttc cac acg cgt ttc cag ccc ggc tac tac      912
Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr
    290                 295                 300

ggc aac gac agc ttc aac tac tgg agc ggc aac tac gtg agc acc cgc      960
Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg
305                 310                 315                 320

ccc agc atc ggc agc aac gac atc atc acc agc ccc ttc tac ggc aac     1008
Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn
                325                 330                 335

aag agc agc gag ccc gtg cag aac ctt gag ttc aac ggc gag aag gtg     1056
Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val
            340                 345                 350

tac cgc gcc gtg gct aac acc aac ctg gcc gtg tgg ccc tct gca gtg     1104
Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val
        355                 360                 365
```

```
tac agc ggc gtg acc aag gtg gag ttc agc cag tac aac gac cag acc      1152
Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr
    370                 375                 380

gac gag gcc agc acc cag acc tac gac agc aag cgc aac gtg ggc gcc      1200
Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala
385                 390                 395                 400

gtg agc tgg gac agc atc gac cag ctg ccc ccc gag acc acc gac gag      1248
Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu
            405                 410                 415

ccc ctg gag aag ggc tac agc cac cag ctg aac tac gtg atg tgc ttc      1296
Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe
        420                 425                 430

ctg atg cag ggc agc cgc ggc acc atc ccc gtg ctg acc tgg acc cac      1344
Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His
    435                 440                 445

aag agc gtc gac ttc ttc aac atg atc gac agc aag aag atc acc cag      1392
Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln
450                 455                 460

ctg ccc ctg gtg aag gcc tac aag ctc cag agc ggc gcc agc gtg gtg      1440
Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val
465                 470                 475                 480

gca ggc ccc cgc ttc acc ggc ggc gac atc atc cag tgc acc gag aac      1488
Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn
            485                 490                 495

ggc agc gcc gcc acc atc tac gtg acc ccc gac gtg agc tac agc cag      1536
Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln
        500                 505                 510

aag tac cgc gcc cgc atc cac tac gcc agc acc agc cag atc acc ttc      1584
Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe
    515                 520                 525

acc ctg agc ctg gac ggg gcc ccc ttc aac caa tac tac ttc gac aag      1632
Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp Lys
530                 535                 540

acc atc aac aag ggc gac acc ctg acc tac aac agc ttc aac ctg gcc      1680
Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala
545                 550                 555                 560

agc ttc agc acc cct ttc gag ctg agc ggc aac aac ctc cag atc ggc      1728
Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly
            565                 570                 575

gtg acc ggc ctg agc gcc ggc gac aag gtg tac atc gac aag atc gag      1776
Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu
        580                 585                 590

ttc atc ccc gtg aac tag atctgagct                                   1803
Phe Ile Pro Val Asn
    595
```

<210> 4

<211> 597

<212> PRT

<213> Artificial Sequence

<400> 4

Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

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

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn
100 105 110

Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser
115 120 125

His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val
130 135 140

Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu
145 150 155 160

Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu
165 170 175

Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr
180 185 190

Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg
195 200 205

Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu

IOT-080-12+PCT.ST25

Modified Cry3A Toxins.ST25

210 215 220

Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp
225 230 235 240

Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val
245 250 255

Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr
260 265 270

Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp
275 280 285

Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr
290 295 300

Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg
305 310 315 320

Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn
325 330 335

Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val
340 345 350

Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val
355 360 365

Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr
370 375 380

Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala
385 390 395 400

Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu
405 410 415

Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe
420 425 430

Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His
435 440 445

Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln
450 455 460

tgttcctgct gaaggacgcc caaatctacg gagaggagtg gggctacgag aaggaggaca 540
tcgccgagtt ctacaagcgc cagctgaagc tgacccagga gtacaccgac cactgcgtga 600
agtggtacaa cgtgggtcta gacaagctcc gcggcagcag ctacgagagc tgggtgaact 660
tcaaccgcta ccgccgcgag atgaccctga ccgtgctgga cctgatcgcc ctgttccccc 720
tgtacgacgt gcgcctgtac cccaaggagg tgaagaccga gctgacccgc gacgtgctga 780
ccgaccccat cgtgggcgtg aacaacctgc gcggctacgg caccaccttc agcaacatcg 840
agaactacat ccgcaagccc cacctgttcg actacctgca ccgcatccag ttccacacgc 900
gtttccagcc cggctactac ggcaacgaca gcttcaacta ctggagcggc aactacgtga 960
gcacccgccc cagcatcggc agcaacgaca tcatcaccag ccccttctac ggcaacaaga 1020
gcagcgagcc cgtgcagaac cttgagttca acggcgagaa ggtgtaccgc gccgtggcta 1080
acaccaacct ggccgtgtgg ccctctgcag tgtacagcgg cgtgaccaag gtggagttca 1140
gccagtacaa cgaccagacc gacgaggcca gcacccagac ctacgacagc aagcgcaacg 1200
tgggcgccgt gagctgggac agcatcgacc agctgccccc cgagaccacc gacgagcccc 1260
tggagaaggg ctacagccac cagctgaact acgtgatgtg cttcctgatg cagggcagcc 1320
gcggcaccat ccccgtgctg acctggaccc acaagagcgt cgacttcttc aacatgatcg 1380
acagcaagaa gatcacccag ctgccccctgg tgaaggccta caagctccag agcggcgcca 1440
gcgtggtggc aggccccgc ttcaccggcg gcgacatcat ccagtgcacc gagaacggca 1500
gcgccgccac catctacgtg accccgacg tgagctacag ccagaagtac cgcgcccgca 1560
tccactacgc cagcaccagc cagatcacct tcaccctgag cctggacggg gccccttca 1620
accaatacta cttcgacaag accatcaaca agggcgacac cctgacctac aacagcttca 1680
acctggccag cttcagcacc cctttcgagc tgagcggcaa caacctccag atcggcgtga 1740
ccggcctgag cgccggcgac aaggtgtaca tcgacaagat cgagttcatc cccgtgaact 1800
agatctgagc tcaagatctg ttgtacaaaa accagcaact cactgcactg cacttcactt 1860
cacttcactg tatgaataaa agtctggtgt ctggttcctg atcgatgact gactactcca 1920
ctttgtgcag aacttagtat gtatttgtat ttgtaaaata cttctatcaa taaaatttct 1980
aattcctaaa accaaaatcc agtgggtacc gaattcactg gccgtcgttt tacaacgtcg 2040
tgactgggaa aaccctggcg ttacccaact taatcgcctt gcagcacatc cccctttcgc 2100
cagctggcgt aatagcgaag aggcccgcac cgatcgccct tcccaacagt tgcgcagcct 2160
gaatggcgaa tggcgcctga tgcggtattt tctccttacg catctgtgcg gtatttcaca 2220
ccgcatatgg tgcactctca gtacaatctg ctctgatgcc gcatagttaa gccagccccg 2280
acacccgcca acacccgctg acgcgccctg acgggcttgt ctgctcccgg catccgctta 2340
cagacaagct gtgaccgtct ccgggagctg catgtgtcag aggttttcac cgtcatcacc 2400

gaaacgcgcg agacgaaagg gcctcgtgat acgcctattt ttataggtta atgtcatgat 2460
aataatggtt tcttagacgt caggtggcac ttttcgggga aatgtgcgcg gaacccctat 2520
ttgtttattt ttctaaatac attcaaatat gtatccgctc atgagacaat aaccctgata 2580
aatgcttcaa taatattgaa aaaggaagag tatgagtatt caacatttcc gtgtcgccct 2640
tattcccttt tttgcggcat tttgccttcc tgtttttgct cacccagaaa cgctggtgaa 2700
agtaaaagat gctgaagatc agttgggtgc acgagtgggt tacatcgaac tggatctcaa 2760
cagcggtaag atccttgaga gttttcgccc cgaagaacgt tttccaatga tgagcacttt 2820
taaagttctg ctatgtggcg cggtattatc ccgtattgac gccgggcaag agcaactcgg 2880
tcgccgcata cactattctc agaatgactt ggttgagtac tcaccagtca cagaaaagca 2940
tcttacggat ggcatgacag taagagaatt atgcagtgct gccataacca tgagtgataa 3000
cactgcggcc aacttacttc tgacaacgat cggaggaccg aaggagctaa ccgctttttt 3060
gcacaacatg ggggatcatg taactcgcct tgatcgttgg gaaccggagc tgaatgaagc 3120
cataccaaac gacgagcgtg acaccacgat gcctgtagca atggcaacaa cgttgcgcaa 3180
actattaact ggcgaactac ttactctagc ttcccggcaa caattaatag actggatgga 3240
ggcggataaa gttgcaggac cacttctgcg ctcggccctt ccggctggct ggtttattgc 3300
tgataaatct ggagccggtg agcgtgggtc tcgcggtatc attgcagcac tggggccaga 3360
tggtaagccc tcccgtatcg tagttatcta cacgacgggg agtcaggcaa ctatggatga 3420
acgaaataga cagatcgctg agataggtgc ctcactgatt aagcattggt aactgtcaga 3480
ccaagtttac tcatatatac tttagattga tttaaaactt catttttaat ttaaaaggat 3540
ctaggtgaag atccttttg ataatctcat gaccaaaatc ccttaacgtg agttttcgtt 3600
ccactgagcg tcagaccccg tagaaaagat caaaggatct tcttgagatc ctttttttct 3660
gcgcgtaatc tgctgcttgc aaacaaaaaa accaccgcta ccagcggtgg tttgtttgcc 3720
ggatcaagag ctaccaactc tttttccgaa ggtaactggc ttcagcagag cgcagatacc 3780
aaatactgtc cttctagtgt agccgtagtt aggccaccac ttcaagaact ctgtagcacc 3840
gcctacatac ctcgctctgc taatcctgtt accagtggct gctgccagtg gcgataagtc 3900
gtgtcttacc gggttggact caagacgata gttaccggat aaggcgcagc ggtcgggctg 3960
aacgggggt tcgtgcacac agcccagctt ggagcgaacg acctacaccg aactgagata 4020
cctacagcgt gagctatgag aaagcgccac gcttcccgaa gggagaaagg cggacaggta 4080
tccggtaagc ggcagggtcg gaacaggaga gcgcacgagg gagcttccag ggggaaacgc 4140
ctggtatctt tatagtcctg tcgggtttcg ccacctctga cttgagcgtc gatttttgtg 4200
atgctcgtca gggggggcgga gcctatggaa aaacgccagc aacgcggcct ttttacggtt 4260

cctggccttt tgctggcctt ttgctcacat gttctttcct gcgttatccc ctgattctgt 4320
ggataaccgt attaccgcct ttgagtgagc tgataccgct cgccgcagcc gaacgaccga 4380
gtgcagcgag tcagtgagcg aggaagcgga agagcgccca atacgcaaac cgcctctccc 4440
cgcgcgttgg ccgattcatt aatgcagctg gcacgacagg tttcccgact ggaaagcggg 4500
cagtgagcgc aacgcaatta atgtgagtta gctcactcat taggcacccc aggctttaca 4560
ctttatgctt ccggctcgta tgttgtgtgg aattgtgagc ggataacaat ttcacacagg 4620
aaacagctat gaccatgatt acgccaagct tgcacatgac aacaattgta agaggatgga 4680
gaccacaacg atccaacaat acttctgcga cgggctgtga agtatagaga agttaaacgc 4740
ccaaaagcca ttgtgtttgg aatttttagt tattctattt ttcatgatgt atcttcctct 4800
aacatgcctt aatttgcaaa tttggtataa ctactgattg aaaatatatg tatgtaaaaa 4860
aatactaagc atatttgtga agctaaacat gatgttattt aagaaaatat gttgttaaca 4920
tgaataagatt aatatcgaaa tggaaacatc tgtaaattag aatcatctta caagctaaga 4980
gatgttcacg ctttgagaaa cttcttcaga tcatgaccgt agaagtagct ctccaagact 5040
caacgaaggc tgctgcaatt ccacaaatgc atgacatgca tccttgtaac cgtcgtcgcc 5100
gctataaaca cggataactc aattccctgc tccatcaatt tagaaatgag caagcaagca 5160
tccgatcgct caccccatat gcaccaatct gactcccaag tctctgtttc gcattagtac 5220
cgccagcact ccacctatag ctaccaattg agacctttcc agcctaagca gatcgattga 5280
tcgttagagt caaagagttg gtggtacggg tactttaact accatggaat gatggggcgt 5340
tgatgtagagc ggaaagcgcc tccctacgcg gaacaacacc ctcgccatgc cgctcgacta 5400
cagcctcctc ctcgtcggcc gcccacaacg agggagcccg tggtcgcagc caccgaccag 5460
tcatgtctctg tgtcctcgtc cgacctcgac atgtcatggc aaacagtcgg acgccagcac 5520
cagactgacg acatgagtct ctgaagagcc cgccacctag aaagatccga gccctgctgc 5580
tggtagtggt aaccattttc gtcgcgctga cgcggagagc gagaggccag aaatttatag 5640
cgactgacgc tgtggcaggc acgctatcgg aggttacgac gtggcgggtc actcgacgcg 5700
gagttcacag gtcctatcct tgcatcgctc gggccggagt ttacgggact tatccttacg 5760
acgtgctcta aggttgcgat aacgggcgga ggaaggcgtg tggcgtgcgg agacggttta 5820
tacacgtagt gtgcgggagt gtgtttcgta gacgcgggaa agcacgacga cttacgaagg 5880
ttagtggagg aggaggacac actaaaatca ggacgcaaga aactcttcta ttatagtagt 5940
agagaagaga ttataggagt gtgggttgat tctaaagaaa atcgacgcag gacaaccgtc 6000
aaaacgggtg ctttaatata gtagatatat atatatagag agagagagaa agtacaaagg 6060
atgcatttgt gtctgcatat gatcggagta ttactaacgg ccgtcgtaag aaggtccatc 6120
atgcgtggag cgagcccatt tggttggttg tcaggccgca gttaaggcct ccatatatga 6180

ttgtcgtcgg gcccataaca gcatctcctc caccagttta ttgtaagaat aaattaagta 6240
gagatatttg tcgtcgggca gaagaaactt ggacaagaag aagaagcaag ctaggccaat 6300
ttcttgccgg caagaggaag atagtggcct ctagtttata tatcggcgtg atgatgatgc 6360
tcctagctag aaatgagaga agaaaaacgg acgcgtgttt ggtgtgtgtc aatggcgtcc 6420
atccttccat cagatcagaa cgatgaaaaa gtcaagcacg gcatgcatag tatatgtata 6480
gcttgtttta gtgtggcttt gctgagacga atgaaagcaa cggcgggcat atttttcagt 6540
ggctgtagct ttcaggctga aagagacgtg gcatgcaata attcagggaa ttcgtcagcc 6600
aattgaggta gctagtcaac ttgtacattg gtgcgagcaa ttttccgcac tcaggagggc 6660
tagtttgaga gtccaaaaac tataggagat taaagaggct aaaatcctct ccttatttaa 6720
ttttaaataa gtagtgtatt tgtattttaa ctcctccaac ccttccgatt ttatggctct 6780
caaactagca ttcagtctaa tgcatgcatg cttggctaga ggtcgtatgg ggttgttaat 6840
agcatagcta gctacaagtt aaccgggtct tttatattta ataaggacag gcaaagtatt 6900
acttacaaat aaagaataaa gctaggacga actcgtggat tattactaaa tcgaaatgga 6960
cgtaatattc caggcaagaa taattgttcg atcaggagac aagtggggca ttggaccggt 7020
ccttgcaagc aagagcctat ggcgtggtga cacggcgcgt tgcccataca tcatgcctcc 7080
atcgatgatc catcctcact tgctataaaa agaggtgtcc atggtgctca agctcagcca 7140
agcaaataag acgacttgtt tcattgattc ttcaagagat cgagcttctt ttgcaccaca 7200
aggtcgag 7208

<210> 6
<211> 1801
<212> DNA
<213> Artificial Sequence

<220>
<221> CDS
<222> (1)..(1791)
<223> Maize optimized modified cry3A054 coding sequence.

<220>
<221> misc_feature
<222> (322)..(333)

<223> cathepsin G recognition site coding sequence

<400> 6

```
atg acg gcc gac aac aac acc gag gcc ctg gac agc agc acc acc aag      48
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1               5                  10                  15

gac gtg atc cag aag ggc atc agc gtg gtg ggc gac ctg ctg ggc gtg      96
Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
            20                  25                  30

gtg ggc ttc ccc ttc ggc ggc gcc ctg gtg agc ttc tac acc aac ttc     144
Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
        35                  40                  45

ctg aac acc atc tgg ccc agc gag gac ccc tgg aag gcc ttc atg gag     192
Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
    50                  55                  60

cag gtg gag gcc ctg atg gac cag aag atc gcc gac tac gcc aag aac     240
Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65                  70                  75                  80

aag gca ctg gcc gag cta cag ggc ctc cag aac aac gtg gag gac tat     288
Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
                85                  90                  95

gtg agc gcc ctg agc agc tgg cag aag aac ccc gct gca ccg ttc ccc     336
Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Pro
            100                 105                 110

cac agc cag ggc cgc atc cgc gag ctg ttc agc cag gcc gag agc cac     384
His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser His
        115                 120                 125

ttc cgc aac agc atg ccc agc ttc gcc atc agc ggc tac gag gtg ctg     432
Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val Leu
    130                 135                 140

ttc ctg acc acc tac gcc cag gcc gcc aac acc cac ctg ttc ctg ctg     480
Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu Leu
145                 150                 155                 160

aag gac gcc caa atc tac gga gag gag tgg ggc tac gag aag gag gac     528
Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu Asp
                165                 170                 175

atc gcc gag ttc tac aag cgc cag ctg aag ctg acc cag gag tac acc     576
Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr Thr
            180                 185                 190

gac cac tgc gtg aag tgg tac aac gtg ggt cta gac aag ctc cgc ggc     624
Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg Gly
        195                 200                 205

agc agc tac gag agc tgg gtg aac ttc aac cgc tac cgc cgc gag atg     672
Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu Met
    210                 215                 220

acc ctg acc gtg ctg gac ctg atc gcc ctg ttc ccc ctg tac gac gtg     720
```

```
Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp Val
225                 230                 235                 240

cgc ctg tac ccc aag gag gtg aag acc gag ctg acc cgc gac gtg ctg      768
Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val Leu
                245                 250                 255

acc gac ccc atc gtg ggc gtg aac aac ctg cgc ggc tac ggc acc acc      816
Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr Thr
            260                 265                 270

ttc agc aac atc gag aac tac atc cgc aag ccc cac ctg ttc gac tac      864
Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp Tyr
        275                 280                 285

ctg cac cgc atc cag ttc cac acg cgt ttc cag ccc ggc tac tac ggc      912
Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr Gly
    290                 295                 300

aac gac agc ttc aac tac tgg agc ggc aac tac gtg agc acc cgc ccc      960
Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg Pro
305                 310                 315                 320

agc atc ggc agc aac gac atc atc acc agc ccc ttc tac ggc aac aag     1008
Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn Lys
                325                 330                 335

agc agc gag ccc gtg cag aac ctt gag ttc aac ggc gag aag gtg tac     1056
Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val Tyr
            340                 345                 350

cgc gcc gtg gct aac acc aac ctg gcc gtg tgg ccc tct gca gtg tac     1104
Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val Tyr
        355                 360                 365

agc ggc gtg acc aag gtg gag ttc agc cag tac aac gac cag acc gac     1152
Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr Asp
    370                 375                 380

gag gcc agc acc cag acc tac gac agc aag cgc aac gtg ggc gcc gtg     1200
Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala Val
385                 390                 395                 400

agc tgg gac agc atc gac cag ctg ccc ccc gag acc acc gac gag ccc     1248
Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu Pro
                405                 410                 415

ctg gag aag ggc tac agc cac cag ctg aac tac gtg atg tgc ttc ctg     1296
Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe Leu
            420                 425                 430

atg cag ggc agc cgc ggc acc atc ccc gtg ctg acc tgg acc cac aag     1344
Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His Lys
        435                 440                 445

agc gtc gac ttc ttc aac atg atc gac agc aag aag atc acc cag ctg     1392
Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln Leu
    450                 455                 460

ccc ctg gtg aag gcc tac aag ctc cag agc ggc gcc agc gtg gtg gca     1440
Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val Ala
465                 470                 475                 480
```

```
ggc ccc cgc ttc acc ggc ggc gac atc atc cag tgc acc gag aac ggc      1488
Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn Gly
            485                 490                 495

agc gcc gcc acc atc tac gtg acc ccc gac gtg agc tac agc cag aag      1536
Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln Lys
        500                 505                 510

tac cgc gcc cgc atc cac tac gcc agc acc agc cag atc acc ttc acc      1584
Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe Thr
    515                 520                 525

ctg agc ctg gac ggg gcc ccc ttc aac caa tac tac ttc gac aag acc      1632
Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp Lys Thr
530                 535                 540

atc aac aag ggc gac acc ctg acc tac aac agc ttc aac ctg gcc agc      1680
Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala Ser
545                 550                 555                 560

ttc agc acc cct ttc gag ctg agc ggc aac aac ctc cag atc ggc gtg      1728
Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly Val
            565                 570                 575

acc ggc ctg agc gcc ggc gac aag gtg tac atc gac aag atc gag ttc      1776
Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu Phe
        580                 585                 590

atc ccc gtg aac tag atctgagctc                                      1801
Ile Pro Val Asn
    595
```

<210> 7

<211> 596

<212> PRT

<213> Artificial Sequence

<220>

<221> misc_feature

<222> (322)..(333)

<223> cathepsin G recognition site coding sequence

<400> 7

```
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1               5                   10                  15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
            20                  25                  30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
        35                  40                  45
```

290 295 300

Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg Pro
305 310 315 320

Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn Lys
325 330 335

Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val Tyr
340 345 350

Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val Tyr
355 360 365

Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr Asp
370 375 380

Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala Val
385 390 395 400

Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu Pro
405 410 415

Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe Leu
420 425 430

Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His Lys
435 440 445

Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln Leu
450 455 460

Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val Ala
465 470 475 480

Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn Gly
485 490 495

Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln Lys
500 505 510

Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe Thr
515 520 525

Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp Lys Thr
530 535 540

Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala Ser
545 550 555 560

Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly Val
565 570 575

Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu Phe
580 585 590

Ile Pro Val Asn
595

<210> 8

<211> 1807

<212> DNA

<213> Artificial Sequence

<220>

<221> CDS

<222> (1)..(1806)

<223> Maize optimized modified cry3A055 coding sequence.

<220>

<221> misc_feature

<222> (322)..(333)

<223> Cthepsin G recognition site coding sequence.

<400> 8

atg acg gcc gac aac aac acc gag gcc ctg gac agc agc acc acc aag 48
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

gac gtg atc cag aag ggc atc agc gtg gtg ggc gac ctg ctg ggc gtg 96
Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

gtg ggc ttc ccc ttc ggc ggc gcc ctg gtg agc ttc tac acc aac ttc 144
Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

ctg aac acc atc tgg ccc agc gag gac ccc tgg aag gcc ttc atg gag 192
Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

```
cag gtg gag gcc ctg atg gac cag aag atc gcc gac tac gcc aag aac      240
Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65                  70                  75                  80

aag gca ctg gcc gag cta cag ggc ctc cag aac aac gtg gag gac tat      288
Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
                85                  90                  95

gtg agc gcc ctg agc agc tgg cag aag aac ccc gct gca ccg ttc cgc      336
Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Arg
            100                 105                 110

aac ccc cac agc cag ggc cgc atc cgc gag ctg ttc agc cag gcc gag      384
Asn Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu
        115                 120                 125

agc cac ttc cgc aac agc atg ccc agc ttc gcc atc agc ggc tac gag      432
Ser His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu
    130                 135                 140

gtg ctg ttc ctg acc acc tac gcc cag gcc gcc aac acc cac ctg ttc      480
Val Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe
145                 150                 155                 160

ctg ctg aag gac gcc caa atc tac gga gag gag tgg ggc tac gag aag      528
Leu Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys
                165                 170                 175

gag gac atc gcc gag ttc tac aag cgc cag ctg aag ctg acc cag gag      576
Glu Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu
            180                 185                 190

tac acc gac cac tgc gtg aag tgg tac aac gtg ggt cta gac aag ctc      624
Tyr Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu
        195                 200                 205

cgc ggc agc agc tac gag agc tgg gtg aac ttc aac cgc tac cgc cgc      672
Arg Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg
    210                 215                 220

gag atg acc ctg acc gtg ctg gac ctg atc gcc ctg ttc ccc ctg tac      720
Glu Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr
225                 230                 235                 240

gac gtg cgc ctg tac ccc aag gag gtg aag acc gag ctg acc cgc gac      768
Asp Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp
                245                 250                 255

gtg ctg acc gac ccc atc gtg ggc gtg aac aac ctg cgc ggc tac ggc      816
Val Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly
            260                 265                 270

acc acc ttc agc aac atc gag aac tac atc cgc aag ccc cac ctg ttc      864
Thr Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe
        275                 280                 285

gac tac ctg cac cgc atc cag ttc cac acg cgt ttc cag ccc ggc tac      912
Asp Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr
    290                 295                 300

tac ggc aac gac agc ttc aac tac tgg agc ggc aac tac gtg agc acc      960
Tyr Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr
305                 310                 315                 320
```

565 570 575

ggc gtg acc ggc ctg agc gcc ggc gac aag gtg tac atc gac aag atc 1776
Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile
580 585 590

gag ttc atc ccc gtg aac tag atc tga gct c 1807
Glu Phe Ile Pro Val Asn Ile Ala
595 600

<210> 9

<211> 598

<212> PRT

<213> Artificial Sequence

<220>

<221> misc_feature

<222> (322)..(333)

<223> Cthepsin G recognition site coding sequence.

<400> 9

Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

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

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Arg
100 105 110

Asn Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu
115 120 125

Thr Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly
385 390 395 400

Ala Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp
405 410 415

Glu Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys
420 425 430

Phe Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr
435 440 445

His Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr
450 455 460

Gln Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val
465 470 475 480

Val Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu
485 490 495

Asn Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser
500 505 510

Gln Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr
515 520 525

Phe Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Tyr Phe Asp
530 535 540

Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu
545 550 555 560

Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile
565 570 575

Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile
580 585 590

Glu Phe Ile Pro Val Asn
595

<210> 10

<211> 1818

<212> DNA

<213> Artificial Sequence

<220>

<221> CDS

<222> (1)..(1818)

<223> Maize optimized modified cry3A085 coding sequence.

<220>

<221> misc_feature

<222> (346)..(357)

<223> Cathepsin G recognition site coding sequence.

<400> 10

```
atg aac tac aag gag ttc ctc cgc atg acc gcc gac aac aac acc gag      48
Met Asn Tyr Lys Glu Phe Leu Arg Met Thr Ala Asp Asn Asn Thr Glu
 1               5                  10                  15

gcc ctg gac agc agc acc acc aag gac gtg atc cag aag ggc atc agc      96
Ala Leu Asp Ser Ser Thr Thr Lys Asp Val Ile Gln Lys Gly Ile Ser
            20                  25                  30

gtg gtg ggc gac ctg ctg ggc gtg gtg ggc ttc ccc ttc ggc ggc gcc     144
Val Val Gly Asp Leu Leu Gly Val Val Gly Phe Pro Phe Gly Gly Ala
        35                  40                  45

ctg gtg agc ttc tac acc aac ttc ctg aac acc atc tgg ccc agc gag     192
Leu Val Ser Phe Tyr Thr Asn Phe Leu Asn Thr Ile Trp Pro Ser Glu
    50                  55                  60

gac ccc tgg aag gcc ttc atg gag cag gtg gag gcc ctg atg gac cag     240
Asp Pro Trp Lys Ala Phe Met Glu Gln Val Glu Ala Leu Met Asp Gln
65                  70                  75                  80

aag atc gcc gac tac gcc aag aac aag gca ctg gcc gag cta cag ggc     288
Lys Ile Ala Asp Tyr Ala Lys Asn Lys Ala Leu Ala Glu Leu Gln Gly
                85                  90                  95

ctc cag aac aac gtg gag gac tat gtg agc gcc ctg agc agc tgg cag     336
Leu Gln Asn Asn Val Glu Asp Tyr Val Ser Ala Leu Ser Ser Trp Gln
            100                 105                 110

aag aac ccc gct gca ccg ttc cgc aac ccc cac agc cag ggc cgc atc     384
Lys Asn Pro Ala Ala Pro Phe Arg Asn Pro His Ser Gln Gly Arg Ile
        115                 120                 125

cgc gag ctg ttc agc cag gcc gag agc cac ttc cgc aac agc atg ccc     432
Arg Glu Leu Phe Ser Gln Ala Glu Ser His Phe Arg Asn Ser Met Pro
    130                 135                 140

agc ttc gcc atc agc ggc tac gag gtg ctg ttc ctg acc acc tac gcc     480
Ser Phe Ala Ile Ser Gly Tyr Glu Val Leu Phe Leu Thr Thr Tyr Ala
```

Modified Cry3A Toxins.ST25

```
145                 150                 155                 160

cag gcc gcc aac acc cac ctg ttc ctg ctg aag gac gcc caa atc tac      528
Gln Ala Ala Asn Thr His Leu Phe Leu Leu Lys Asp Ala Gln Ile Tyr
                165                 170                 175

gga gag gag tgg ggc tac gag aag gag gac atc gcc gag ttc tac aag      576
Gly Glu Glu Trp Gly Tyr Glu Lys Glu Asp Ile Ala Glu Phe Tyr Lys
            180                 185                 190

cgc cag ctg aag ctg acc cag gag tac acc gac cac tgc gtg aag tgg      624
Arg Gln Leu Lys Leu Thr Gln Glu Tyr Thr Asp His Cys Val Lys Trp
        195                 200                 205

tac aac gtg ggt cta gac aag ctc cgc ggc agc agc tac gag agc tgg      672
Tyr Asn Val Gly Leu Asp Lys Leu Arg Gly Ser Ser Tyr Glu Ser Trp
    210                 215                 220

gtg aac ttc aac cgc tac cgc cgc gag atg acc ctg acc gtg ctg gac      720
Val Asn Phe Asn Arg Tyr Arg Arg Glu Met Thr Leu Thr Val Leu Asp
225                 230                 235                 240

ctg atc gcc ctg ttc ccc ctg tac gac gtg cgc ctg tac ccc aag gag      768
Leu Ile Ala Leu Phe Pro Leu Tyr Asp Val Arg Leu Tyr Pro Lys Glu
                245                 250                 255

gtg aag acc gag ctg acc cgc gac gtg ctg acc gac ccc atc gtg ggc      816
Val Lys Thr Glu Leu Thr Arg Asp Val Leu Thr Asp Pro Ile Val Gly
            260                 265                 270

gtg aac aac ctg cgc ggc tac ggc acc acc ttc agc aac atc gag aac      864
Val Asn Asn Leu Arg Gly Tyr Gly Thr Thr Phe Ser Asn Ile Glu Asn
        275                 280                 285

tac atc cgc aag ccc cac ctg ttc gac tac ctg cac cgc atc cag ttc      912
Tyr Ile Arg Lys Pro His Leu Phe Asp Tyr Leu His Arg Ile Gln Phe
    290                 295                 300

cac acg cgt ttc cag ccc ggc tac tac ggc aac gac agc ttc aac tac      960
His Thr Arg Phe Gln Pro Gly Tyr Tyr Gly Asn Asp Ser Phe Asn Tyr
305                 310                 315                 320

tgg agc ggc aac tac gtg agc acc cgc ccc agc atc ggc agc aac gac     1008
Trp Ser Gly Asn Tyr Val Ser Thr Arg Pro Ser Ile Gly Ser Asn Asp
                325                 330                 335

atc atc acc agc ccc ttc tac ggc aac aag agc agc gag ccc gtg cag     1056
Ile Ile Thr Ser Pro Phe Tyr Gly Asn Lys Ser Ser Glu Pro Val Gln
            340                 345                 350

aac ctt gag ttc aac ggc gag aag gtg tac cgc gcc gtg gct aac acc     1104
Asn Leu Glu Phe Asn Gly Glu Lys Val Tyr Arg Ala Val Ala Asn Thr
        355                 360                 365

aac ctg gcc gtg tgg ccc tct gca gtg tac agc ggc gtg acc aag gtg     1152
Asn Leu Ala Val Trp Pro Ser Ala Val Tyr Ser Gly Val Thr Lys Val
    370                 375                 380

gag ttc agc cag tac aac gac cag acc gac gag gcc agc acc cag acc     1200
Glu Phe Ser Gln Tyr Asn Asp Gln Thr Asp Glu Ala Ser Thr Gln Thr
385                 390                 395                 400

tac gac agc aag cgc aac gtg ggc gcc gtg agc tgg gac agc atc gac     1248
```

```
Tyr Asp Ser Lys Arg Asn Val Gly Ala Val Ser Trp Asp Ser Ile Asp
                405                 410                 415

cag ctg ccc ccc gag acc acc gac gag ccc ctg gag aag ggc tac agc      1296
Gln Leu Pro Pro Glu Thr Thr Asp Glu Pro Leu Glu Lys Gly Tyr Ser
            420                 425                 430

cac cag ctg aac tac gtg atg tgc ttc ctg atg cag ggc agc cgc ggc      1344
His Gln Leu Asn Tyr Val Met Cys Phe Leu Met Gln Gly Ser Arg Gly
        435                 440                 445

acc atc ccc gtg ctg acc tgg acc cac aag agc gtc gac ttc ttc aac      1392
Thr Ile Pro Val Leu Thr Trp Thr His Lys Ser Val Asp Phe Phe Asn
    450                 455                 460

atg atc gac agc aag aag atc acc cag ctg ccc ctg gtg aag gcc tac      1440
Met Ile Asp Ser Lys Lys Ile Thr Gln Leu Pro Leu Val Lys Ala Tyr
465                 470                 475                 480

aag ctc cag agc ggc gcc agc gtg gtg gca ggc ccc cgc ttc acc ggc      1488
Lys Leu Gln Ser Gly Ala Ser Val Val Ala Gly Pro Arg Phe Thr Gly
                485                 490                 495

ggc gac atc atc cag tgc acc gag aac ggc agc gcc gcc acc atc tac      1536
Gly Asp Ile Ile Gln Cys Thr Glu Asn Gly Ser Ala Ala Thr Ile Tyr
            500                 505                 510

gtg acc ccc gac gtg agc tac agc cag aag tac cgc gcc cgc atc cac      1584
Val Thr Pro Asp Val Ser Tyr Ser Gln Lys Tyr Arg Ala Arg Ile His
        515                 520                 525

tac gcc agc acc agc cag atc acc ttc acc ctg agc ctg gac ggg gcc      1632
Tyr Ala Ser Thr Ser Gln Ile Thr Phe Thr Leu Ser Leu Asp Gly Ala
    530                 535                 540

ccc ttc aac caa tac tac ttc gac aag acc atc aac aag ggc gac acc      1680
Pro Phe Asn Gln Tyr Tyr Phe Asp Lys Thr Ile Asn Lys Gly Asp Thr
545                 550                 555                 560

ctg acc tac aac agc ttc aac ctg gcc agc ttc agc acc cct ttc gag      1728
Leu Thr Tyr Asn Ser Phe Asn Leu Ala Ser Phe Ser Thr Pro Phe Glu
                565                 570                 575

ctg agc ggc aac aac ctc cag atc ggc gtg acc ggc ctg agc gcc ggc      1776
Leu Ser Gly Asn Asn Leu Gln Ile Gly Val Thr Gly Leu Ser Ala Gly
            580                 585                 590

gac aag gtg tac atc gac aag atc gag ttc atc ccc gtg aac               1818
Asp Lys Val Tyr Ile Asp Lys Ile Glu Phe Ile Pro Val Asn
        595                 600                 605
```

<210> 11

<211> 606

<212> PRT

<213> Artificial Sequence

<220>

WO 2013/081840 PCT/US2012/063101

<221> misc_feature

<222> (346)..(357)

<223> Cathepsin G recognition site coding sequence.

<400> 11

```
Met Asn Tyr Lys Glu Phe Leu Arg Met Thr Ala Asp Asn Asn Thr Glu
1               5                  10                  15

Ala Leu Asp Ser Ser Thr Thr Lys Asp Val Ile Gln Lys Gly Ile Ser
            20                  25                  30

Val Val Gly Asp Leu Leu Gly Val Val Gly Phe Pro Phe Gly Gly Ala
        35                  40                  45

Leu Val Ser Phe Tyr Thr Asn Phe Leu Asn Thr Ile Trp Pro Ser Glu
    50                  55                  60

Asp Pro Trp Lys Ala Phe Met Glu Gln Val Glu Ala Leu Met Asp Gln
65                  70                  75                  80

Lys Ile Ala Asp Tyr Ala Lys Asn Lys Ala Leu Ala Glu Leu Gln Gly
                85                  90                  95

Leu Gln Asn Asn Val Glu Asp Tyr Val Ser Ala Leu Ser Ser Trp Gln
            100                 105                 110

Lys Asn Pro Ala Ala Pro Phe Arg Asn Pro His Ser Gln Gly Arg Ile
        115                 120                 125

Arg Glu Leu Phe Ser Gln Ala Glu Ser His Phe Arg Asn Ser Met Pro
    130                 135                 140

Ser Phe Ala Ile Ser Gly Tyr Glu Val Leu Phe Leu Thr Thr Tyr Ala
145                 150                 155                 160

Gln Ala Ala Asn Thr His Leu Phe Leu Leu Lys Asp Ala Gln Ile Tyr
                165                 170                 175

Gly Glu Glu Trp Gly Tyr Glu Lys Glu Asp Ile Ala Glu Phe Tyr Lys
            180                 185                 190

Arg Gln Leu Lys Leu Thr Gln Glu Tyr Thr Asp His Cys Val Lys Trp
        195                 200                 205

Tyr Asn Val Gly Leu Asp Lys Leu Arg Gly Ser Ser Tyr Glu Ser Trp
    210                 215                 220
```

Val Asn Phe Asn Arg Tyr Arg Arg Glu Met Thr Leu Thr Val Leu Asp
225 230 235 240

Leu Ile Ala Leu Phe Pro Leu Tyr Asp Val Arg Leu Tyr Pro Lys Glu
245 250 255

Val Lys Thr Glu Leu Thr Arg Asp Val Leu Thr Asp Pro Ile Val Gly
260 265 270

Val Asn Asn Leu Arg Gly Tyr Gly Thr Thr Phe Ser Asn Ile Glu Asn
275 280 285

Tyr Ile Arg Lys Pro His Leu Phe Asp Tyr Leu His Arg Ile Gln Phe
290 295 300

His Thr Arg Phe Gln Pro Gly Tyr Tyr Gly Asn Asp Ser Phe Asn Tyr
305 310 315 320

Trp Ser Gly Asn Tyr Val Ser Thr Arg Pro Ser Ile Gly Ser Asn Asp
325 330 335

Ile Ile Thr Ser Pro Phe Tyr Gly Asn Lys Ser Ser Glu Pro Val Gln
340 345 350

Asn Leu Glu Phe Asn Gly Glu Lys Val Tyr Arg Ala Val Ala Asn Thr
355 360 365

Asn Leu Ala Val Trp Pro Ser Ala Val Tyr Ser Gly Val Thr Lys Val
370 375 380

Glu Phe Ser Gln Tyr Asn Asp Gln Thr Asp Glu Ala Ser Thr Gln Thr
385 390 395 400

Tyr Asp Ser Lys Arg Asn Val Gly Ala Val Ser Trp Asp Ser Ile Asp
405 410 415

Gln Leu Pro Pro Glu Thr Thr Asp Glu Pro Leu Glu Lys Gly Tyr Ser
420 425 430

His Gln Leu Asn Tyr Val Met Cys Phe Leu Met Gln Gly Ser Arg Gly
435 440 445

Thr Ile Pro Val Leu Thr Trp Thr His Lys Ser Val Asp Phe Phe Asn
450 455 460

Met Ile Asp Ser Lys Lys Ile Thr Gln Leu Pro Leu Val Lys Ala Tyr

465 470 475 480

Lys Leu Gln Ser Gly Ala Ser Val Val Ala Gly Pro Arg Phe Thr Gly
485 490 495

Gly Asp Ile Ile Gln Cys Thr Glu Asn Gly Ser Ala Ala Thr Ile Tyr
500 505 510

Val Thr Pro Asp Val Ser Tyr Ser Gln Lys Tyr Arg Ala Arg Ile His
515 520 525

Tyr Ala Ser Thr Ser Gln Ile Thr Phe Thr Leu Ser Leu Asp Gly Ala
530 535 540

Pro Phe Asn Gln Tyr Tyr Phe Asp Lys Thr Ile Asn Lys Gly Asp Thr
545 550 555 560

Leu Thr Tyr Asn Ser Phe Asn Leu Ala Ser Phe Ser Thr Pro Phe Glu
565 570 575

Leu Ser Gly Asn Asn Leu Gln Ile Gly Val Thr Gly Leu Ser Ala Gly
580 585 590

Asp Lys Val Tyr Ile Asp Lys Ile Glu Phe Ile Pro Val Asn
595 600 605

<210> 12

<211> 1794

<212> DNA

<213> Artificial Sequence

<220>

<221> CDS

<222> (1)..(1794)

<223> Maize optimized modified cry3A082 coding sequence.

<220>

<221> misc_feature

<222> (1609)..(1620)

<223> Cathepsin G recognition site coding sequence

<400> 12

```
atg acg gcc gac aac aac acc gag gcc ctg gac agc agc acc acc aag        48
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1               5                  10                  15

gac gtg atc cag aag ggc atc agc gtg gtg ggc gac ctg ctg ggc gtg        96
Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
            20                  25                  30

gtg ggc ttc ccc ttc ggc ggc gcc ctg gtg agc ttc tac acc aac ttc       144
Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
        35                  40                  45

ctg aac acc atc tgg ccc agc gag gac ccc tgg aag gcc ttc atg gag       192
Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
    50                  55                  60

cag gtg gag gcc ctg atg gac cag aag atc gcc gac tac gcc aag aac       240
Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65                  70                  75                  80

aag gca ctg gcc gag cta cag ggc ctc cag aac aac gtg gag gac tat       288
Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
                85                  90                  95

gtg agc gcc ctg agc agc tgg cag aag aac ccc gtc tcg agc cgc aac       336
Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn
            100                 105                 110

ccc cac agc cag ggc cgc atc cgc gag ctg ttc agc cag gcc gag agc       384
Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser
        115                 120                 125

cac ttc cgc aac agc atg ccc agc ttc gcc atc agc ggc tac gag gtg       432
His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val
    130                 135                 140

ctg ttc ctg acc acc tac gcc cag gcc gcc aac acc cac ctg ttc ctg       480
Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu
145                 150                 155                 160

ctg aag gac gcc caa atc tac gga gag gag tgg ggc tac gag aag gag       528
Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu
                165                 170                 175

gac atc gcc gag ttc tac aag cgc cag ctg aag ctg acc cag gag tac       576
Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr
            180                 185                 190

acc gac cac tgc gtg aag tgg tac aac gtg ggt cta gac aag ctc cgc       624
Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg
        195                 200                 205

ggc agc agc tac gag agc tgg gtg aac ttc aac cgc tac cgc cgc gag       672
Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu
    210                 215                 220

atg acc ctg acc gtg ctg gac ctg atc gcc ctg ttc ccc ctg tac gac       720
Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp
225                 230                 235                 240
```

```
gtg cgc ctg tac ccc aag gag gtg aag acc gag ctg acc cgc gac gtg      768
Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val
                245                 250                 255

ctg acc gac ccc atc gtg ggc gtg aac aac ctg cgc ggc tac ggc acc      816
Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr
            260                 265                 270

acc ttc agc aac atc gag aac tac atc cgc aag ccc cac ctg ttc gac      864
Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp
        275                 280                 285

tac ctg cac cgc atc cag ttc cac acg cgt ttc cag ccc ggc tac tac      912
Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr
    290                 295                 300

ggc aac gac agc ttc aac tac tgg agc ggc aac tac gtg agc acc cgc      960
Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg
305                 310                 315                 320

ccc agc atc ggc agc aac gac atc atc acc agc ccc ttc tac ggc aac     1008
Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn
                325                 330                 335

aag agc agc gag ccc gtg cag aac ctt gag ttc aac ggc gag aag gtg     1056
Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val
            340                 345                 350

tac cgc gcc gtg gct aac acc aac ctg gcc gtg tgg ccc tct gca gtg     1104
Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val
        355                 360                 365

tac agc ggc gtg acc aag gtg gag ttc agc cag tac aac gac cag acc     1152
Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr
    370                 375                 380

gac gag gcc agc acc cag acc tac gac agc aag cgc aac gtg ggc gcc     1200
Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala
385                 390                 395                 400

gtg agc tgg gac agc atc gac cag ctg ccc ccc gag acc acc gac gag     1248
Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu
                405                 410                 415

ccc ctg gag aag ggc tac agc cac cag ctg aac tac gtg atg tgc ttc     1296
Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe
            420                 425                 430

ctg atg cag ggc agc cgc ggc acc atc ccc gtg ctg acc tgg acc cac     1344
Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His
        435                 440                 445

aag agc gtc gac ttc ttc aac atg atc gac agc aag aag atc acc cag     1392
Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln
    450                 455                 460

ctg ccc ctg gtg aag gcc tac aag ctc cag agc ggc gcc agc gtg gtg     1440
Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val
465                 470                 475                 480

gca ggc ccc cgc ttc acc ggc ggc gac atc atc cag tgc acc gag aac     1488
Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn
                485                 490                 495
```

```
ggc agc gcc gcc acc atc tac gtg acc ccc gac gtg agc tac agc cag      1536
Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln
            500                 505                 510

aag tac cgc gcc cgc atc cac tac gcc agc acc agc cag atc acc ttc      1584
Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe
        515                 520                 525

acc ctg agc ctg gac ggg gcc ccc gct gca ccg ttc tac ttc gac aag      1632
Thr Leu Ser Leu Asp Gly Ala Pro Ala Ala Pro Phe Tyr Phe Asp Lys
    530                 535                 540

acc atc aac aag ggc gac acc ctg acc tac aac agc ttc aac ctg gcc      1680
Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala
545                 550                 555                 560

agc ttc agc acc cct ttc gag ctg agc ggc aac aac ctc cag atc ggc      1728
Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly
                565                 570                 575

gtg acc ggc ctg agc gcc ggc gac aag gtg tac atc gac aag atc gag      1776
Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu
            580                 585                 590

ttc atc ccc gtg aac tag                                             1794
Phe Ile Pro Val Asn
        595
```

US20130315884A1

<210> 13

<211> 597

<212> PRT

<213> Artificial Sequence

<220>

<221> misc_feature

<222> (1609)..(1620)

<223> Cathepsin G recognition site coding sequence

<400> 13

```
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1               5                   10                  15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
            20                  25                  30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
        35                  40                  45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
```

50 55 60

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

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn
100 105 110

Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser
115 120 125

His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val
130 135 140

Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu
145 150 155 160

Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu
165 170 175

Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr
180 185 190

Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg
195 200 205

Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu
210 215 220

Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp
225 230 235 240

Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val
245 250 255

Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr
260 265 270

Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp
275 280 285

Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr
290 295 300

Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg
305 310 315 320

Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn
325 330 335

Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val
340 345 350

Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val
355 360 365

Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr
370 375 380

Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala
385 390 395 400

Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu
405 410 415

Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe
420 425 430

Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His
435 440 445

Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln
450 455 460

Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val
465 470 475 480

Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn
485 490 495

Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln
500 505 510

Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe
515 520 525

Thr Leu Ser Leu Asp Gly Ala Pro Ala Ala Pro Phe Tyr Phe Asp Lys
530 535 540

Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu Ala
545 550 555 560

Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile Gly
565 570 575

Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile Glu
580 585 590

Phe Ile Pro Val Asn
595

<210> 14

<211> 1816

<212> DNA

<213> Artificial Sequence

<220>

<221> CDS

<222> (1)..(1812)

<223> Maize optimized modified cry3A058 coding sequence.

<220>

<221> misc_feature

<222> (1621)..(1632)

<223> Cathepsin G recognition site coding sequence

<400> 14

atg acg gcc gac aac aac acc gag gcc ctg gac agc agc acc acc aag 48
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

gac gtg atc cag aag ggc atc agc gtg gtg ggc gac ctg ctg ggc gtg 96
Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

gtg ggc ttc ccc ttc ggc ggc gcc ctg gtg agc ttc tac acc aac ttc 144
Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

ctg aac acc atc tgg ccc agc gag gac ccc tgg aag gcc ttc atg gag 192
Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

cag gtg gag gcc ctg atg gac cag aag atc gcc gac tac gcc aag aac 240
Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65 70 75 80

```
aag gca ctg gcc gag cta cag ggc ctc cag aac aac gtg gag gac tat      288
Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
                85                  90                  95

gtg agc gcc ctg agc agc tgg cag aag aac ccc gtc tcg agc cgc aac      336
Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn
           100                 105                 110

ccc cac agc cag ggc cgc atc cgc gag ctg ttc agc cag gcc gag agc      384
Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser
       115                 120                 125

cac ttc cgc aac agc atg ccc agc ttc gcc atc agc ggc tac gag gtg      432
His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val
   130                 135                 140

ctg ttc ctg acc acc tac gcc cag gcc gcc aac acc cac ctg ttc ctg      480
Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu
145                 150                 155                 160

ctg aag gac gcc caa atc tac gga gag gag tgg ggc tac gag aag gag      528
Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu
                165                 170                 175

gac atc gcc gag ttc tac aag cgc cag ctg aag ctg acc cag gag tac      576
Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr
           180                 185                 190

acc gac cac tgc gtg aag tgg tac aac gtg ggt cta gac aag ctc cgc      624
Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg
       195                 200                 205

ggc agc agc tac gag agc tgg gtg aac ttc aac cgc tac cgc cgc gag      672
Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu
   210                 215                 220

atg acc ctg acc gtg ctg gac ctg atc gcc ctg ttc ccc ctg tac gac      720
Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp
225                 230                 235                 240

gtg cgc ctg tac ccc aag gag gtg aag acc gag ctg acc cgc gac gtg      768
Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val
                245                 250                 255

ctg acc gac ccc atc gtg ggc gtg aac aac ctg cgc ggc tac ggc acc      816
Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr
           260                 265                 270

acc ttc agc aac atc gag aac tac atc cgc aag ccc cac ctg ttc gac      864
Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp
       275                 280                 285

tac ctg cac cgc atc cag ttc cac acg cgt ttc cag ccc ggc tac tac      912
Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr
   290                 295                 300

ggc aac gac agc ttc aac tac tgg agc ggc aac tac gtg agc acc cgc      960
Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg
305                 310                 315                 320

ccc agc atc ggc agc aac gac atc atc acc agc ccc ttc tac ggc aac     1008
Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn
```

```
                    325                 330                 335

aag agc agc gag ccc gtg cag aac ctt gag ttc aac ggc gag aag gtg     1056
Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val
            340                 345                 350

tac cgc gcc gtg gct aac acc aac ctg gcc gtg tgg ccc tct gca gtg     1104
Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val
        355                 360                 365

tac agc ggc gtg acc aag gtg gag ttc agc cag tac aac gac cag acc     1152
Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr
    370                 375                 380

gac gag gcc agc acc cag acc tac gac agc aag cgc aac gtg ggc gcc     1200
Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala
385                 390                 395                 400

gtg agc tgg gac agc atc gac cag ctg ccc ccc gag acc acc gac gag     1248
Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu
                405                 410                 415

ccc ctg gag aag ggc tac agc cac cag ctg aac tac gtg atg tgc ttc     1296
Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe
            420                 425                 430

ctg atg cag ggc agc cgc ggc acc atc ccc gtg ctg acc tgg acc cac     1344
Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His
        435                 440                 445

aag agc gtc gac ttc ttc aac atg atc gac agc aag aag atc acc cag     1392
Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln
    450                 455                 460

ctg ccc ctg gtg aag gcc tac aag ctc cag agc ggc gcc agc gtg gtg     1440
Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val
465                 470                 475                 480

gca ggc ccc cgc ttc acc ggc ggc gac atc atc cag tgc acc gag aac     1488
Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn
                485                 490                 495

ggc agc gcc gcc acc atc tac gtg acc ccc gac gtg agc tac agc cag     1536
Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln
            500                 505                 510

aag tac cgc gcc cgc atc cac tac gcc agc acc agc cag atc acc ttc     1584
Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe
        515                 520                 525

acc ctg agc ctg gac ggg gcc ccc ttc aac caa tac gct gca ccg ttc     1632
Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Ala Ala Pro Phe
    530                 535                 540

tac ttc gac aag acc atc aac aag ggc gac acc ctg acc tac aac agc     1680
Tyr Phe Asp Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser
545                 550                 555                 560

ttc aac ctg gcc agc ttc agc acc cct ttc gag ctg agc ggc aac aac     1728
Phe Asn Leu Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn
                565                 570                 575

ctc cag atc ggc gtg acc ggc ctg agc gcc ggc gac aag gtg tac atc     1776
```

Leu Gln Ile Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile
580 585 590

gac aag atc gag ttc atc ccc gtg aac tag atc tga gctc 1816
Asp Lys Ile Glu Phe Ile Pro Val Asn Ile
595 600

<210> 15

<211> 601

<212> PRT

<213> Artificial Sequence

<220>

<221> misc_feature

<222> (1621)..(1632)

<223> Cathepsin G recognition site coding sequence

<400> 15

Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

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

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Val Ser Ser Arg Asn
100 105 110

Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser
115 120 125

His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val
130 135 140

WO 2005/121310 PCT/US2005/020401

Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu
145 150 155 160

Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu
165 170 175

Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr
180 185 190

Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg
195 200 205

Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu
210 215 220

Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp
225 230 235 240

Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val
245 250 255

Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr
260 265 270

Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp
275 280 285

Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr
290 295 300

Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg
305 310 315 320

Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn
325 330 335

Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val
340 345 350

Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val
355 360 365

Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr
370 375 380

Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala
385 390 395 400

Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu
405 410 415

Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe
420 425 430

Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His
435 440 445

Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln
450 455 460

Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val
465 470 475 480

Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn
485 490 495

Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln
500 505 510

Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe
515 520 525

Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Ala Ala Pro Phe
530 535 540

Tyr Phe Asp Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser
545 550 555 560

Phe Asn Leu Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn
565 570 575

Leu Gln Ile Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile
580 585 590

Asp Lys Ile Glu Phe Ile Pro Val Asn
595 600

<210> 16

<211> 1813

<212> DNA

<213> Artificial Sequence

<220>

<221> CDS

<222> (1)..(1812)

<223> Maize optimized modified cry3A057 coding sequence.

<220>

<221> misc_feature

<222> (322)..(333)

<223> Cathepsin G recognition site coding sequence

<220>

<221> misc_feature

<222> (1618)..(1629)

<223> Cathepsin G recognition site coding sequence

<400> 16

```
atg acg gcc gac aac aac acc gag gcc ctg gac agc agc acc acc aag      48
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
 1               5                  10                  15

gac gtg atc cag aag ggc atc agc gtg gtg ggc gac ctg ctg ggc gtg      96
Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
            20                  25                  30

gtg ggc ttc ccc ttc ggc ggc gcc ctg gtg agc ttc tac acc aac ttc     144
Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
        35                  40                  45

ctg aac acc atc tgg ccc agc gag gac ccc tgg aag gcc ttc atg gag     192
Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
    50                  55                  60

cag gtg gag gcc ctg atg gac cag aag atc gcc gac tac gcc aag aac     240
Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65                  70                  75                  80

aag gca ctg gcc gag cta cag ggc ctc cag aac aac gtg gag gac tat     288
Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
                85                  90                  95

gtg agc gcc ctg agc agc tgg cag aag aac ccc gct gca ccg ttc ccc     336
Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Pro
            100                 105                 110

cac agc cag ggc cgc atc cgc gag ctg ttc agc cag gcc gag agc cac     384
His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser His
        115                 120                 125
```

```
ttc cgc aac agc atg ccc agc ttc gcc atc agc ggc tac gag gtg ctg      432
Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val Leu
    130                 135                 140

ttc ctg acc acc tac gcc cag gcc gcc aac acc cac ctg ttc ctg ctg      480
Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu Leu
145                 150                 155                 160

aag gac gcc caa atc tac gga gag gag tgg ggc tac gag aag gag gac      528
Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu Asp
                165                 170                 175

atc gcc gag ttc tac aag cgc cag ctg aag ctg acc cag gag tac acc      576
Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr Thr
            180                 185                 190

gac cac tgc gtg aag tgg tac aac gtg ggt cta gac aag ctc cgc ggc      624
Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg Gly
        195                 200                 205

agc agc tac gag agc tgg gtg aac ttc aac cgc tac cgc cgc gag atg      672
Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu Met
    210                 215                 220

acc ctg acc gtg ctg gac ctg atc gcc ctg ttc ccc ctg tac gac gtg      720
Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp Val
225                 230                 235                 240

cgt ctg tac ccc aag gag gtg aag acc gag ctg acc cgc gac gtg ctg      768
Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val Leu
                245                 250                 255

acc gac ccc atc gtg ggc gtg aac aac ctg cgc ggc tac ggc acc acc      816
Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr Thr
            260                 265                 270

ttc agc aac atc gag aac tac atc cgc aag ccc cac ctg ttc gac tac      864
Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp Tyr
        275                 280                 285

ctg cac cgc atc cag ttc cac acg cgt ttc cag ccc ggc tac tac ggc      912
Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr Gly
    290                 295                 300

aac gac agc ttc aac tac tgg agc ggc aac tac gtg agc acc cgc ccc      960
Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg Pro
305                 310                 315                 320

agc atc ggc agc aac gac atc atc acc agc ccc ttc tac ggc aac aag     1008
Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn Lys
                325                 330                 335

agc agc gag ccc gtg cag aac ctt gag ttc aac ggc gag aag gtg tac     1056
Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val Tyr
            340                 345                 350

cgc gcc gtg gct aac acc aac ctg gcc gtg tgg ccc tct gca gtg tac     1104
Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val Tyr
        355                 360                 365

agc ggc gtg acc aag gtg gag ttc agc cag tac aac gac cag acc gac     1152
Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr Asp
```

```
    370                 375                 380

gag gcc agc acc cag acc tac gac agc aag cgc aac gtg ggc gcc gtg      1200
Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala Val
385                 390                 395                 400

agc tgg gac agc atc gac cag ctg ccc ccc gag acc acc gac gag ccc      1248
Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu Pro
                405                 410                 415

ctg gag aag ggc tac agc cac cag ctg aac tac gtg atg tgc ttc ctg      1296
Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe Leu
            420                 425                 430

atg cag ggc agc cgc ggc acc atc ccc gtg ctg acc tgg acc cac aag      1344
Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His Lys
        435                 440                 445

agc gtc gac ttc ttc aac atg atc gac agc aag aag atc acc cag ctg      1392
Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln Leu
    450                 455                 460

ccc ctg gtg aag gcc tac aag ctc cag agc ggc gcc agc gtg gtg gca      1440
Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val Ala
465                 470                 475                 480

ggc ccc cgc ttc acc ggc ggc gac atc atc cag tgc acc gag aac ggc      1488
Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn Gly
                485                 490                 495

agc gcc gcc acc atc tac gtg acc ccc gac gtg agc tac agc cag aag      1536
Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln Lys
            500                 505                 510

tac cgc gcc cgc atc cac tac gcc agc acc agc cag atc acc ttc acc      1584
Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe Thr
        515                 520                 525

ctg agc ctg gac ggg gcc ccc ttc aac caa tac gct gca ccg ttc tac      1632
Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Ala Ala Pro Phe Tyr
    530                 535                 540

ttc gac aag acc atc aac aag ggc gac acc ctg acc tac aac agc ttc      1680
Phe Asp Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe
545                 550                 555                 560

aac ctg gcc agc ttc agc acc cct ttc gag ctg agc ggc aac aac ctc      1728
Asn Leu Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu
                565                 570                 575

cag atc ggc gtg acc ggc ctg agc gcc ggc gac aag gtg tac atc gac      1776
Gln Ile Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp
            580                 585                 590

aag atc gag ttc atc ccc gtg aac tag atc tga gct c                    1813
Lys Ile Glu Phe Ile Pro Val Asn     Ile     Ala
        595                 600
```

<210> 17

<211> 600

<212> PRT

<213> Artificial Sequence

<220>

<221> misc_feature

<222> (322)..(333)

<223> Cathepsin G recognition site coding sequence

<220>

<221> misc_feature

<222> (1618)..(1629)

<223> Cathepsin G recognition site coding sequence

<400> 17

Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

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

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Pro
100 105 110

His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu Ser His
115 120 125

Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu Val Leu
130 135 140

Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe Leu Leu
145 150 155 160

Modified Cry3A Toxins.ST25

Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys Glu Asp
165 170 175

Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu Tyr Thr
180 185 190

Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu Arg Gly
195 200 205

Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg Glu Met
210 215 220

Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr Asp Val
225 230 235 240

Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp Val Leu
245 250 255

Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly Thr Thr
260 265 270

Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe Asp Tyr
275 280 285

Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr Tyr Gly
290 295 300

Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr Arg Pro
305 310 315 320

Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly Asn Lys
325 330 335

Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys Val Tyr
340 345 350

Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala Val Tyr
355 360 365

Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln Thr Asp
370 375 380

Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly Ala Val
385 390 395 400

Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp Glu Pro
405 410 415

TOP SECRET

Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys Phe Leu
420 425 430

Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr His Lys
435 440 445

Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr Gln Leu
450 455 460

Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val Val Ala
465 470 475 480

Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu Asn Gly
485 490 495

Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser Gln Lys
500 505 510

Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr Phe Thr
515 520 525

Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Ala Ala Pro Phe Tyr
530 535 540

Phe Asp Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe
545 550 555 560

Asn Leu Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu
565 570 575

Gln Ile Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp
580 585 590

Lys Ile Glu Phe Ile Pro Val Asn
595 600

<210> 18

<211> 1819

<212> DNA

<213> Artificial Sequence

<220>

<221> CDS

10/30/2014 05:15:21 PM

<222> (1)..(1818)

<223> Maize optimized modified cry3A056 coding sequence.

<220>

<221> misc_feature

<222> (322)..(333)

<223> Catthepsin G recognition site coding sequence.

<220>

<221> misc_feature

<222> (1624)..(1635)

<223> Catthepsin G recognition site coding sequence.

<400> 18

```
atg acg gcc gac aac aac acc gag gcc ctg gac agc agc acc acc aag        48
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
 1               5                  10                  15

gac gtg atc cag aag ggc atc agc gtg gtg ggc gac ctg ctg ggc gtg        96
Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
            20                  25                  30

gtg ggc ttc ccc ttc ggc ggc gcc ctg gtg agc ttc tac acc aac ttc       144
Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
        35                  40                  45

ctg aac acc atc tgg ccc agc gag gac ccc tgg aag gcc ttc atg gag       192
Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
    50                  55                  60

cag gtg gag gcc ctg atg gac cag aag atc gcc gac tac gcc aag aac       240
Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65                  70                  75                  80

aag gca ctg gcc gag cta cag ggc ctc cag aac aac gtg gag gac tat       288
Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
                85                  90                  95

gtg agc gcc ctg agc agc tgg cag aag aac ccc gct gca ccg ttc cgc       336
Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Arg
            100                 105                 110

aac ccc cac agc cag ggc cgc atc cgc gag ctg ttc agc cag gcc gag       384
Asn Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu
        115                 120                 125

agc cac ttc cgc aac agc atg ccc agc ttc gcc atc agc ggc tac gag       432
Ser His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu
    130                 135                 140
```

```
gtg ctg ttc ctg acc acc tac gcc cag gcc gcc aac acc cac ctg ttc      480
Val Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe
145                 150                 155                 160

ctg ctg aag gac gcc caa atc tac gga gag gag tgg ggc tac gag aag      528
Leu Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys
                165                 170                 175

gag gac atc gcc gag ttc tac aag cgc cag ctg aag ctg acc cag gag      576
Glu Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu
            180                 185                 190

tac acc gac cac tgc gtg aag tgg tac aac gtg ggt cta gac aag ctc      624
Tyr Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu
        195                 200                 205

cgc ggc agc agc tac gag agc tgg gtg aac ttc aac cgc tac cgc cgc      672
Arg Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg
    210                 215                 220

gag atg acc ctg acc gtg ctg gac ctg atc gcc ctg ttc ccc ctg tac      720
Glu Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr
225                 230                 235                 240

gac gtg cgc ctg tac ccc aag gag gtg aag acc gag ctg acc cgc gac      768
Asp Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp
                245                 250                 255

ttg ctg acc gac ccc atc gtg ggc gtg aac aac ctg cgc ggc tac ggc      816
Leu Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly
            260                 265                 270

acc acc ttc agc aac atc gag aac tac atc cgc aag ccc cac ctg ttc      864
Thr Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe
        275                 280                 285

gac tac ctg cac cgc atc cag ttc cac acg cgt ttc cag ccc ggc tac      912
Asp Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr
    290                 295                 300

tac ggc aac gac agc ttc aac tac tgg agc ggc aac tac gtg agc acc      960
Tyr Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr
305                 310                 315                 320

cgc ccc agc atc ggc agc aac gac atc atc acc agc ccc ttc tac ggc     1008
Arg Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly
                325                 330                 335

aac aag agc agc gag ccc gtg cag aac ctt gag ttc aac ggc gag aag     1056
Asn Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys
            340                 345                 350

gtg tac cgc gcc gtg gct aac acc aac ctg gcc gtg tgg ccc tct gca     1104
Val Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala
        355                 360                 365

gtg tac agc ggc gtg acc aag gtg gag ttc agc cag tac aac gac cag     1152
Val Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln
    370                 375                 380

acc gac gag gcc agc acc cag acc tac gac agc aag cgc aac gtg ggc     1200
Thr Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly
```

```
385                 390                 395                 400

gcc gtg agc tgg gac agc atc gac cag ctg ccc ccc gag acc acc gac     1248
Ala Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp
                405                 410                 415

gag ccc ctg gag aag ggc tac agc cac cag ctg aac tac gtg atg tgc     1296
Glu Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys
            420                 425                 430

ttc ctg atg cag ggc agc cgc ggc acc atc ccc gtg ctg acc tgg acc     1344
Phe Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr
        435                 440                 445

cac aag agc gtc gac ttc ttc aac atg atc gac agc aag aag atc acc     1392
His Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr
    450                 455                 460

cag ctg ccc ctg gtg aag gcc tac aag ctc cag agc ggc gcc agc gtg     1440
Gln Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val
465                 470                 475                 480

gtg gca ggc ccc cgc ttc acc ggc ggc gac atc atc cag tgc acc gag     1488
Val Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu
                485                 490                 495

aac ggc agc gcc gcc acc atc tac gtg acc ccc gac gtg agc tac agc     1536
Asn Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser
            500                 505                 510

cag aag tac cgc gcc cgc atc cac tac gcc agc acc agc cag atc acc     1584
Gln Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr
        515                 520                 525

ttc acc ctg agc ctg gac ggg gcc ccc ttc aac caa tac gct gca ccg     1632
Phe Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Ala Ala Pro
    530                 535                 540

ttc tac ttc gac aag acc atc aac aag ggc gac acc ctg acc tac aac     1680
Phe Tyr Phe Asp Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn
545                 550                 555                 560

agc ttc aac ctg gcc agc ttc agc acc cct ttc gag ctg agc ggc aac     1728
Ser Phe Asn Leu Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn
                565                 570                 575

aac ctc cag atc ggc gtg acc ggc ctg agc gcc ggc gac aag gtg tac     1776
Asn Leu Gln Ile Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr
            580                 585                 590

atc gac aag atc gag ttc atc ccc gtg aac tag atc tga gct c           1819
Ile Asp Lys Ile Glu Phe Ile Pro Val Asn     Ile     Ala
        595                 600
```

<210> 19

<211> 602

<212> PRT

<213> Artificial Sequence

<220>

<221> misc_feature

<222> (322)..(333)

<223> Catthepsin G recognition site coding sequence.

<220>

<221> misc_feature

<222> (1624)..(1635)

<223> Catthepsin G recognition site coding sequence.

<400> 19

Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

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

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Arg
100 105 110

Asn Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu
115 120 125

Ser His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu
130 135 140

Val Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe
145 150 155 160

Leu Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys
165 170 175

Glu Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu
180 185 190

Tyr Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu
195 200 205

Arg Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg
210 215 220

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

Asp Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp
245 250 255

Val Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly
260 265 270

Thr Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe
275 280 285

Asp Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr
290 295 300

Tyr Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr
305 310 315 320

Arg Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly
325 330 335

Asn Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys
340 345 350

Val Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala
355 360 365

Val Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln
370 375 380

Thr Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly
385 390 395 400

Ala Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp
405 410 415

Glu Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys
420 425 430

Phe Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr
435 440 445

His Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr
450 455 460

Gln Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val
465 470 475 480

Val Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu
485 490 495

Asn Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser
500 505 510

Gln Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr
515 520 525

Phe Thr Leu Ser Leu Asp Gly Ala Pro Phe Asn Gln Tyr Ala Ala Pro
530 535 540

Phe Tyr Phe Asp Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn
545 550 555 560

Ser Phe Asn Leu Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn
565 570 575

Asn Leu Gln Ile Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr
580 585 590

Ile Asp Lys Ile Glu Phe Ile Pro Val Asn
595 600

<210> 20

<211> 1797

<212> DNA

<213> Artificial Sequence

<220>

<221> CDS

<222> (1)..(1791)

<223> Maize optimized modified cry3A083 coding sequence.

IDT 083101.2197.ST25

<220>

<221> misc_feature

<222> (322)..(333)

<223> Cathepsin G recognition site coding sequence.

<220>

<221> misc_feature

<222> (1612)..(1623)

<223> cathepsin G recognition site coding sequence

<400> 20

atg acg gcc gac aac aac acc gag gcc ctg gac agc agc acc acc aag
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1 5 10 15

gac gtg atc cag aag ggc atc agc gtg gtg ggc gac ctg ctg ggc gtg
Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
20 25 30

gtg ggc ttc ccc ttc ggc ggc gcc ctg gtg agc ttc tac acc aac ttc
Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
35 40 45

ctg aac acc atc tgg ccc agc gag gac ccc tgg aag gcc ttc atg gag
Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
50 55 60

cag gtg gag gcc ctg atg gac cag aag atc gcc gac tac gcc aag aac
Gln Val Glu Ala Leu Met Asp Gln Lys Ile Ala Asp Tyr Ala Lys Asn
65 70 75 80

aag gca ctg gcc gag cta cag ggc ctc cag aac aac gtg gag gac tat
Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
85 90 95

gtg agc gcc ctg agc agc tgg cag aag aac ccc gct gca ccg ttc cgc
Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Arg
100 105 110

aac ccc cac agc cag ggc cgc atc cgc gag ctg ttc agc cag gcc gag
Asn Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu
115 120 125

agc cac ttc cgc aac agc atg ccc agc ttc gcc atc agc ggc tac gag
Ser His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu
130 135 140

gtg ctg ttc ctg acc acc tac gcc cag gcc gcc aac acc cac ctg ttc
Val Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe
145 150 155 160

```
ctg ctg aag gac gcc caa atc tac gga gag gag tgg ggc tac gag aag      528
Leu Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys
                165                 170                 175

gag gac atc gcc gag ttc tac aag cgc cag ctg aag ctg acc cag gag      576
Glu Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu
            180                 185                 190

tac acc gac cac tgc gtg aag tgg tac aac gtg ggt cta gac aag ctc      624
Tyr Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu
        195                 200                 205

cgc ggc agc agc tac gag agc tgg gtg aac ttc aac cgc tac cgc cgc      672
Arg Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg
    210                 215                 220

gag atg acc ctg acc gtg ctg gac ctg atc gcc ctg ttc ccc ctg tac      720
Glu Met Thr Leu Thr Val Leu Asp Leu Ile Ala Leu Phe Pro Leu Tyr
225                 230                 235                 240

gac gtg cgc ctg tac ccc aag gag gtg aag acc gag ctg acc cgc gac      768
Asp Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp
                245                 250                 255

gtg ctg acc gac ccc atc gtg ggc gtg aac aac ctg cgc ggc tac ggc      816
Val Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly
            260                 265                 270

acc acc ttc agc aac atc gag aac tac atc cgc aag ccc cac ctg ttc      864
Thr Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe
        275                 280                 285

gac tac ctg cac cgc atc cag ttc cac acg cgt ttc cag ccc ggc tac      912
Asp Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr
    290                 295                 300

tac ggc aac gac agc ttc aac tac tgg agc ggc aac tac gtg agc acc      960
Tyr Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr
305                 310                 315                 320

cgc ccc agc atc ggc agc aac gac atc atc acc agc ccc ttc tac ggc     1008
Arg Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly
                325                 330                 335

aac aag agc agc gag ccc gtg cag aac ctt gag ttc aac ggc gag aag     1056
Asn Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys
            340                 345                 350

gtg tac cgc gcc gtg gct aac acc aac ctg gcc gtg tgg ccc tct gca     1104
Val Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala
        355                 360                 365

gtg tac agc ggc gtg acc aag gtg gag ttc agc cag tac aac gac cag     1152
Val Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln
    370                 375                 380

acc gac gag gcc agc acc cag acc tac gac agc aag cgc aac gtg ggc     1200
Thr Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly
385                 390                 395                 400

gcc gtg agc tgg gac agc atc gac cag ctg ccc ccc gag acc acc gac     1248
Ala Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp
```

```
                    405                 410                 415

gag ccc ctg gag aag ggc tac agc cac cag ctg aac tac gtg atg tgc     1296
Glu Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys
            420                 425                 430

ttc ctg atg cag ggc agc cgc ggc acc atc ccc gtg ctg acc tgg acc     1344
Phe Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr
        435                 440                 445

cac aag agc gtc gac ttc ttc aac atg atc gac agc aag aag atc acc     1392
His Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr
    450                 455                 460

cag ctg ccc ctg gtg aag gcc tac aag ctc cag agc ggc gcc agc gtg     1440
Gln Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val
465                 470                 475                 480

gtg gca ggc ccc cgc ttc acc ggc ggc gac atc atc cag tgc acc gag     1488
Val Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu
                485                 490                 495

aac ggc agc gcc gcc acc atc tac gtg acc ccc gac gtg agc tac agc     1536
Asn Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser
            500                 505                 510

cag aag tac cgc gcc cgc atc cac tac gcc agc acc agc cag atc acc     1584
Gln Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr
        515                 520                 525

ttc acc ctg agc ctg gac ggg gcc ccc gct gca ccg ttc tac ttc gac     1632
Phe Thr Leu Ser Leu Asp Gly Ala Pro Ala Ala Pro Phe Tyr Phe Asp
    530                 535                 540

aag acc atc aac aag ggc gac acc ctg acc tac aac agc ttc aac ctg     1680
Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu
545                 550                 555                 560

gcc agc ttc agc acc cct ttc gag ctg agc ggc aac aac ctc cag atc     1728
Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile
                565                 570                 575

ggc gtg acc ggc ctg agc gcc ggc gac aag gtg tac atc gac aag atc     1776
Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile
            580                 585                 590

gag ttc atc ccc gtg aactag                                         1797
Glu Phe Ile Pro Val
        595
```

<210> 21

<211> 597

<212> PRT

<213> Artificial Sequence

<220>

<221> misc_feature

<222> (322)..(333)

<223> Cathepsin G recognition site coding sequence.

<220>

<221> misc_feature

<222> (1612)..(1623)

<223> cathepsin G recognition site coding sequence

<400> 21

```
Met Thr Ala Asp Asn Asn Thr Glu Ala Leu Asp Ser Ser Thr Thr Lys
1               5                  10                  15

Asp Val Ile Gln Lys Gly Ile Ser Val Val Gly Asp Leu Leu Gly Val
            20                  25                  30

Val Gly Phe Pro Phe Gly Gly Ala Leu Val Ser Phe Tyr Thr Asn Phe
        35                  40                  45

Leu Asn Thr Ile Trp Pro Ser Glu Asp Pro Trp Lys Ala Phe Met Glu
    50                  55                  60

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

Lys Ala Leu Ala Glu Leu Gln Gly Leu Gln Asn Asn Val Glu Asp Tyr
                85                  90                  95

Val Ser Ala Leu Ser Ser Trp Gln Lys Asn Pro Ala Ala Pro Phe Arg
            100                 105                 110

Asn Pro His Ser Gln Gly Arg Ile Arg Glu Leu Phe Ser Gln Ala Glu
        115                 120                 125

Ser His Phe Arg Asn Ser Met Pro Ser Phe Ala Ile Ser Gly Tyr Glu
    130                 135                 140

Val Leu Phe Leu Thr Thr Tyr Ala Gln Ala Ala Asn Thr His Leu Phe
145                 150                 155                 160

Leu Leu Lys Asp Ala Gln Ile Tyr Gly Glu Glu Trp Gly Tyr Glu Lys
                165                 170                 175

Glu Asp Ile Ala Glu Phe Tyr Lys Arg Gln Leu Lys Leu Thr Gln Glu
            180                 185                 190
```


EDIT_ME_ST25.txt 03/15/2012

Tyr Thr Asp His Cys Val Lys Trp Tyr Asn Val Gly Leu Asp Lys Leu
195 200 205

Arg Gly Ser Ser Tyr Glu Ser Trp Val Asn Phe Asn Arg Tyr Arg Arg
210 215 220

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

Asp Val Arg Leu Tyr Pro Lys Glu Val Lys Thr Glu Leu Thr Arg Asp
245 250 255

Val Leu Thr Asp Pro Ile Val Gly Val Asn Asn Leu Arg Gly Tyr Gly
260 265 270

Thr Thr Phe Ser Asn Ile Glu Asn Tyr Ile Arg Lys Pro His Leu Phe
275 280 285

Asp Tyr Leu His Arg Ile Gln Phe His Thr Arg Phe Gln Pro Gly Tyr
290 295 300

Tyr Gly Asn Asp Ser Phe Asn Tyr Trp Ser Gly Asn Tyr Val Ser Thr
305 310 315 320

Arg Pro Ser Ile Gly Ser Asn Asp Ile Ile Thr Ser Pro Phe Tyr Gly
325 330 335

Asn Lys Ser Ser Glu Pro Val Gln Asn Leu Glu Phe Asn Gly Glu Lys
340 345 350

Val Tyr Arg Ala Val Ala Asn Thr Asn Leu Ala Val Trp Pro Ser Ala
355 360 365

Val Tyr Ser Gly Val Thr Lys Val Glu Phe Ser Gln Tyr Asn Asp Gln
370 375 380

Thr Asp Glu Ala Ser Thr Gln Thr Tyr Asp Ser Lys Arg Asn Val Gly
385 390 395 400

Ala Val Ser Trp Asp Ser Ile Asp Gln Leu Pro Pro Glu Thr Thr Asp
405 410 415

Glu Pro Leu Glu Lys Gly Tyr Ser His Gln Leu Asn Tyr Val Met Cys
420 425 430

Phe Leu Met Gln Gly Ser Arg Gly Thr Ile Pro Val Leu Thr Trp Thr
435 440 445

His Lys Ser Val Asp Phe Phe Asn Met Ile Asp Ser Lys Lys Ile Thr
450 455 460

Gln Leu Pro Leu Val Lys Ala Tyr Lys Leu Gln Ser Gly Ala Ser Val
465 470 475 480

Val Ala Gly Pro Arg Phe Thr Gly Gly Asp Ile Ile Gln Cys Thr Glu
485 490 495

Asn Gly Ser Ala Ala Thr Ile Tyr Val Thr Pro Asp Val Ser Tyr Ser
500 505 510

Gln Lys Tyr Arg Ala Arg Ile His Tyr Ala Ser Thr Ser Gln Ile Thr
515 520 525

Phe Thr Leu Ser Leu Asp Gly Ala Pro Ala Ala Pro Phe Tyr Phe Asp
530 535 540

Lys Thr Ile Asn Lys Gly Asp Thr Leu Thr Tyr Asn Ser Phe Asn Leu
545 550 555 560

Ala Ser Phe Ser Thr Pro Phe Glu Leu Ser Gly Asn Asn Leu Gln Ile
565 570 575

Gly Val Thr Gly Leu Ser Ala Gly Asp Lys Val Tyr Ile Asp Lys Ile
580 585 590

Glu Phe Ile Pro Val
595

<210> 22

<211> 21

<212> DNA

<213> Artificial Sequence

<400> 22
ggatccacca tgacggccga c 21

<210> 23

<211> 29

<212> DNA

<213> Artificial Sequence

SUBSTITUTE SHEET (RULE 26) 08.10.01

<400> 23
gaacggtgca gcggggttct tctgccagc 29

<210> 24

<211> 29

<212> DNA

<213> Artificial Sequence

<400> 24
gctgcaccgt tccccccacag ccagggccg 29

<210> 25

<211> 21

<212> DNA

<213> Artificial Sequence

<400> 25
cctagaccca cgttgtacca c 21

<210> 26

<211> 29

<212> DNA

<213> Artificial Sequence

<400> 26
gctgcaccgt tccgcaaccc ccacagcca 29

<210> 27

<211> 19

<212> DNA

<213> Artificial Sequence

<400> 27
gagcgtcgac ttcttcaac 19

<210> 28

<211> 30

<212> DNA

<213> Artificial Sequence

<400> 28
gaacggtgca gcgtattggt tgaaggggc 30

<210> 29

<211> 30

<212> DNA

<213> Artificial Sequence

<400> 29
gctgcaccgt tctacttcga caagaccatc 30

<210> 30

<211> 21

<212> DNA

<213> Artificial Sequence

<400> 30
gagctcagat ctagttcacg g 21

<210> 31

<211> 32

<212> DNA

<213> Artificial Sequence

<400> 31
cggggccccc gctgcaccgt tctacttcga ca 32

<210> 32

<211> 32

<212> DNA

<213> Artificial Sequence

<400> 32
tgtcgaagta gaacggtgca gcgggggccc cg 32

<210> 33

<211> 48

<212> DNA

<213> Artificial Sequence

<400> 33
ggatccacca tgaactacaa ggagttcctc cgcatgaccg ccgacaac 48

<210> 34

<211> 20

<212> DNA

<213> Artificial Sequence

<400> 34
tctccacctg ctccatgaag 20

<210> 35

<211> 4

<212> PRT

<213> Protease recognition site.

<400> 35

Ala Ala Pro Phe
1

<210> 36

<211> 4

<212> PRT

<213> Protease recognition site.

<400> 36

Ala Ala Pro Met

1

<210> 37

<211> 4

<212> PRT

<213> Protease recognition site.

<400> 37

Ala Val Pro Phe
1

<210> 38

<211> 4

<212> PRT

<213> Protease recognition site.

<400> 38

Pro Phe Leu Phe

603162121.082101

406
408

REQUISITOS NECESARIOS PARA PRESENTACION Y AGILIZACION DEL TRAMITE

	USUARIO	RADICADOR
PATENTE DE INVENCIÓN MODELO DE UTILIDAD	( )	( )
- Indicación de que se solicita la concesión de una patente.	( )	( )
- Datos de identificación del solicitante o de la persona que se presenta la solicitud.	( )	( )
- Descripción de la invención.	( )	( )
- Dibujos de ser estos pertinentes.	( )	( )
- Comprobante de Pago de las tasas establecidas.		
COMPLETA ( )	INCOMPLETA ( )	( )
DISEÑO INDUSTRIAL ( )		
- Indicación de que se solicita el registro de Diseño Industrial	( )	( )
- Datos de identificación del solicitante o de la persona que presenta	( )	( )
- la solicitud.		
- Representación gráfica y fotográfica del Diseño Industrial o muestra del Material que incorpora el diseño.	( )	( )
Comprobante de pago de las tasas establecidas	( )	( )
COMPLETA ( )	INCOMPLETA ( )	( )
ESQUEMA DE TRAZADO ( )		
- Indicación de que solicita el registro de un Esquema de trazado.	( )	( )
- Datos de identificación del solicitante o de la persona que presenta La solicitud.	( )	( )
Representación gráfica del esquema de trazado	( )	( )
Comprobante de pago de las tasas establecidas.	( )	( )
COMPLETA ( )	INCOMPLETA ( )	( )

Firma Responsable:

Fecha:

SUPERINDUSTRIA Y COMERCIO
Radicacion : 04017803 00000000 Folios: 405
Fecha (AMD): 2004-08-27 16:28:15
Trar : 011 PCT-PATENT 379 FASE NACI 411 PRESENTAC
Responsable: 2020 DIVISION DE NUEVAS CREACIONES

467
409

Industria y Comercio
SUPERINTENDENCIA

2020
Bogotá D C

07378

Doctor(a)
MARGARITA CASTELLANOS A
Apoderado
SYNGENTA PARTICIPATIONS AG

Oficio N°

ASUNTO:

Radicación N°	04-17803
Solicitud internacional	PCT/US01/29613
Fecha inicio fase nacional	31/03/2004
Trámite	011
Evento	379
Actuación	416
Folios	001

Como resultado del examen de forma, realizado con fundamento en el artículo 38 de la Decisión 486 de la Comisión de la Comunidad Andina, artículo 27 del Tratado de Cooperación en Materia de Patentes (PCT), regla 51 bis del R glamento y Circular Única de la Superintendencia de Industria y Comercio, se le comunica que:

- La solicitud no contiene copia del documento en que consta la cesión del derecho de patente del inventor al solicitante o a su causante (Art. 26-k) Decisión 486).
- La solicitud no contiene el poder debidamente otorgado, con el lleno de requisitos exigidos por los artículos 63 y subsiguientes del Código de Procedimiento Civil. Adicionalmente deberá presentar el documento que acredita la existencia y representación legal de la sociedad solicitante, cuando éste fuere una persona jurídica.

En cumplimiento de lo dispuesto por el artículo 39 de la Decisión 486, se le requiere para que dentro del término de dos meses siguientes a la fecha de notificación del presente oficio, complete los requisitos anteriormente enunciados. En caso de no dar cumplimiento al presente requerimiento, la solicitud se considerará abandonada y perderá su prelación.

Notifíquese el presente oficio conforme a lo dispuesto en el numeral 5.2, literal e), del capítulo quinto del título primero de la Circular Única N° 10 de 2001.

ALIX CARMENZA CESPEDES DE VERGEL
Jefe de la División de Nuevas Creaciones.

El anterior requerimiento fue notificado por fijación en lista de fecha,
06 AGO. 2004
202044

Sede Centro: Carrera 13 No. 27-00 Pisos 2, 5, 7 y 10
Conmutador: 3 820340 Fax: 3 505220
Sede CAN: Tr. 40A No. 3B-50 3153265 al 69
Fax: 3 505220. Linea 9800-910 165
Web: www.sic.gov.co e-mail: info@sic.gov.co
Bogotá D.C. - Colombia