

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 397 549**

51 Int. Cl.:

C12N 15/32 (2006.01)

C07K 14/325 (2006.01)

C12N 15/62 (2006.01)

C12N 15/63 (2006.01)

C12N 1/21 (2006.01)

A01H 5/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **08.01.2002 E 08075614 (1)**

97 Fecha y número de publicación de la concesión europea: **14.11.2012 EP 1988099**

54 Título: **Proteínas insecticidas de Bacillus thuringiensis**

30 Prioridad:

09.01.2001 US 331355 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

07.03.2013

73 Titular/es:

**BAYER CROPSCIENCE NV (100.0%)
J.E. Mommaertslaan 14
1831 Diegem, BE**

72 Inventor/es:

**ARNAUT, GRETA;
BOETS, ANNEMIE;
VANNESTE, STIJN y
VAN RIE, JEROEN**

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 397 549 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Proteínas insecticidas de *Bacillus thuringiensis*

5 INTRODUCCIÓN

La presente invención se refiere a nuevas secuencias de ácidos nucleicos, particularmente secuencias de ADN, que codifican proteínas insecticidas producidas por cepas de *Bacillus thuringiensis*. Particularmente, se proporcionan nuevas secuencias de ácidos nucleicos, particularmente secuencias de ADN que codifican proteínas designadas Cry2Af, que son útiles para proteger a las plantas frente a daños causados por insectos. También se incluyen en esta memoria microorganismos y plantas transformadas con una secuencia de ácidos nucleicos, particularmente una secuencia de ADN, que codifica al menos una de las proteínas Cry2A recientemente aisladas.

15 ANTECEDENTES DE LA INVENCION

(i) Campo de la invención:

Bacillus thuringiensis (abreviado en esta memoria como "Bt") es bien conocido por su toxicidad específica a plagas de insectos, y ha sido utilizado desde hace casi un siglo para reprimir las plagas de insectos de plantas. En años más recientes, se produjeron plantas transgénicas que expresan proteínas de *Bt* que se encontró que reprimían con éxito los daños causados por insectos en plantas (p. ej. Vaeck et al., 1987, Jansens et al., 1997).

A pesar del aislamiento de un cierto número de genes de *Bt* insecticidas, continúa la búsqueda de nuevos genes que codifiquen proteínas insecticidas. De hecho, se sabe que proteínas de *Bt* insecticidas tienen una gama de insectos diana relativamente estrecha en comparación con insecticidas químicos. También, el hecho de tener múltiples toxinas contra la misma especie de insecto diana permite el uso de proteínas con diferentes modos de acción, de modo que se puede prevenir o retardar el desarrollo de resistencia de los insectos. Y proteínas Bt insecticidas con diferentes secuencias de aminoácidos tienen diferentes niveles de eficacia insecticida frente a insectos específicos, haciendo deseable tener varias proteínas insecticidas diferentes disponibles con el fin de ser capaces de reprimir las plagas de insectos relevantes de diferentes plantas de cultivo.

(ii) Descripción de Técnica Relacionada:

Previamente se identificaron varios tipos de proteínas Cry2A (véase Crickmore et al., 1998, incorporado en esta memoria como referencia).

La proteína Cry2Ae tiene la mayor identidad en la secuencia de aminoácidos con la proteína Cry2Aa1 (Donovan et al., número de acceso a GenBank M31738), pero sigue difiriendo en aproximadamente un 9 por ciento de su secuencia de aminoácidos.

La identidad de secuencia más próxima a la proteína Cry2Af se encontró en la proteína Cry2Ab1 (Widner y Whiteley, número de acceso a GenBank M23724), pero ambas proteínas siguen difiriendo en aproximadamente el 5 por ciento de su secuencia de aminoácidos.

La identidad de secuencia más próxima a la proteína Cry2Ag se encontró en la proteína Cry2Ac1 (Wu et al., número de acceso a GenBank X57252), pero ambas proteínas siguen difiriendo en aproximadamente un 20 por ciento de sus secuencias de aminoácidos.

Proteínas Cry2A adicionalmente conocidas incluyen la proteína Cry2Ad1 (Choi et al., 1999), y otras proteínas Cry2Aa, Cry2Ab y Cry2Ac (Crickmore et al., 1998). Proteínas similares a Cry2A y secuencias de ADN que las codifican también se muestran en la patente de EE.UU. 5.338.544, en la solicitud de patente PCT publicada WO 00/26371 y en la solicitud de patente PCT publicada WO 98/40490.

La expresión de proteínas de tipo Cry2A en plantas ha sido descrita, p. ej., en Kota et al. (1999) y en la solicitud de patente PCT publicada WO 00/26371.

50 OBJETOS Y SUMARIO DE LA INVENCION

De acuerdo con esta invención se proporciona una secuencia de ácidos nucleicos, particularmente una secuencia de ADN, que codifica una proteína que comprende la secuencia de aminoácidos de la proteína codificada por el gen *cry2Af* depositado en la BCCM-LMBP bajo el número de acceso LMBP 4247.

Particularmente preferida de acuerdo con esta invención es una secuencia de ácidos nucleicos, particularmente una secuencia de ADN, que codifica una proteína que comprende la secuencia de aminoácidos de la proteína de

SEQ ID N° 4

Además, de acuerdo con esta invención se proporcionan secuencias de ácidos nucleicos, particularmente secuencias de ADN que codifican una proteína que comprende la secuencia de aminoácidos de SEQ ID N° 4 desde la posición de aminoácido 1 a la posición de aminoácido 625, o una proteína que comprende la secuencia de aminoácidos de una proteína insecticida codificada por un ADN que se hibrida a la secuencia codificadora de *cry2Af* depositada en la BCCM-LMBP bajo el número de acceso LMBP 4247 bajo condiciones de hibridación rigurosas, teniendo dicho ADN una identidad de la secuencia de al menos el 98% con la secuencia codificante de SEQ ID N° 3.

Además, de acuerdo con esta invención se proporcionan las secuencias de ácidos nucleicos anteriores, particularmente secuencias de ADN, que comprenden una secuencia artificial que tiene una utilización del codón diferente en comparación con la secuencia que se produce de forma natural, pero que codifican la misma proteína o su fragmento insecticida, preferiblemente dicha utilización del codón se asemeja a la de plantas, particularmente la planta hospedante en la que se ha de transformar la secuencia de ácidos nucleicos, particularmente el ADN.

Incluso adicionalmente se proporciona, de acuerdo con esta invención, una proteína que comprende la secuencia de aminoácidos de la proteína codificada por el gen *cry2Af* depositado en la BCCM-LMBP bajo el número de acceso LMBP 4247, una proteína que comprende la secuencia de aminoácidos de SEQ ID N° 4 desde la posición de aminoácido 1 a la posición de aminoácido 625, o una proteína con 5 a 10 aminoácidos añadidos, reemplazados o suprimidos sin modificar la actividad insecticida de la proteína, o una proteína que comprende la secuencia de aminoácidos de una proteína insecticida codificada por un ADN que se hibrida a la secuencia codificadora de *cry2Af* depositada en la BCCM-LMBP bajo el número de acceso LMBP 4247 bajo condiciones de hibridación rigurosas, teniendo dicha proteína una identidad de la secuencia de al menos el 98% de la proteína de la SEQ ID N° 4, de la secuencia codificante de SEQ ID N° 3.

Particularmente preferida en esta memoria es la proteína según se define en el párrafo precedente que comprende la secuencia de aminoácidos de SEQ ID N° 4, o una proteína con menos de 5 aminoácidos añadidos, reemplazados o suprimidos sin modificar la actividad insecticida de la proteína.

También se proporcionan en esta memoria genes quiméricos que comprenden el ADN según se define anteriormente bajo el control de un promotor expresable en plantas, y células vegetales, plantas o semillas transformadas de modo que contengan esos genes quiméricos, particularmente células vegetales, plantas o semillas seleccionadas del grupo que consiste en: maíz, algodón, arroz, tabaco, colza oleaginosa, especies de Brassica, berenjena, soja, patata, girasol, tomate, caña de azúcar, té, habas, tabaco, fresa, clavo, pepino, sandía, pimiento, avena, cebada, trigo, dalia, gladiolo, crisantemo, remolacha azucarera, sorgo, alfalfa y cacahuete. De acuerdo con esta invención, el gen quimérico se puede integrar en el ADN nuclear, del plástido o mitocondrial de las células vegetales, o también puede contener un ADN que codifique un péptido diana o de tránsito eficaz para fijarse como objetivo a la vacuola, cloroplasto, mitocondria, plástido, o para la secreción.

Además, de acuerdo con esta invención se proporcionan microorganismos transformados de modo que contengan cualquiera de las secuencias de ADN anteriores, particularmente los seleccionados del género *Pseudomonas*, *Agrobacterium*, *Escherichia* o *Bacillus*.

También se proporciona en esta memoria un procedimiento para hacer a una planta resistente a un insecto, que comprende transformar células vegetales con los genes quiméricos anteriores, y regenerar plantas transformadas procedentes de este tipo de células que son resistentes a insectos. En una realización, dicho insecto se selecciona del grupo que consiste en: *Heliothis virescens*, *Helicoverpa zea*, *Helicoverpa armígera*, *Anticarsia gemmatalis* y *Ostrinia nubilalis*, *Chilo suppressalis*, *Chilo partellus*, *Scirpophaga incertulas*, *Sesamia inferens*, *Cnaphalocrocis medinalis*, *Marasmia patnalis*, *Marasmia exigua*, *Marasmia ruralis* y *Scirpophaga innotata*.

Esta invención se refiere también al uso de la proteína definida anteriormente para reprimir insectos seleccionados del grupo que consiste en: *Heliothis virescens*, *Helicoverpa zea*, *Helicoverpa armígera*, *Anticarsia gemmatalis* y *Ostrinia nubilalis*, *Chilo suppressalis*, *Chilo partellus*, *Scirpophaga incertulas*, *Sesamia inferens*, *Cnaphalocrocis medinalis*, *Marasmia patnalis*, *Marasmia exigua*, *Marasmia ruralis* y *Scirpophaga innotata*.

Esta descripción describe adicionalmente un método para reprimir plagas de insectos del arroz lepidópteros, particularmente lepidópteros barrenadores del tallo del arroz, patrones de arroz, gusanos cortadores del arroz,

gusanos cogolleros del arroz o enrolladores de la hoja del arroz, preferiblemente un insecto seleccionado del grupo que consiste en: *Chilo suppressalis*, *Chilo partellus*, *Scirpophaga incertulas*, *Sesamia inferens*, *Cnaphalocrocis medinalis*, *Marasmia patnalis*, *Marasmia exigua*, *Marasmia ruralis*, *Scirpophaga innotata*, método que comprende aplicar a una zona o planta a proteger una proteína Cry2A según se define en esta memoria (es decir, plantando una planta de arroz transformada con un gen *cry2A* de esta invención o atomizando una composición que contiene una proteína Cry2A de esta invención). La invención se refiere también al uso anterior de las proteínas de esta invención contra plagas de insectos lepidópteros del arroz para minimizar el daño a las plantas de arroz.

DESCRIPCIÓN DETALLADA DE REALIZACIONES PREFERIDAS DE LA INVENCION

De acuerdo con esta invención, una "secuencia de ácidos nucleicos" se refiere a una molécula de ADN o ARN en forma de cadena sencilla o doble, preferiblemente un ADN o ARN, particularmente un ADN que codifica cualquiera de las proteínas Cry2A de esta invención. Una "secuencia de ácidos nucleicos aislada", tal como se utiliza en esta memoria, se refiere a una secuencia de ácidos nucleicos que ya no se encuentra en el entorno natural del que se aisló, p. ej. a partir de la secuencia de ácidos nucleicos en otro hospedante bacteriano o en el genoma nuclear de la planta.

De acuerdo con esta invención, los términos "proteína" o "polipéptido" se utilizan de forma indistinta para aludir a una secuencia de aminoácidos, sin referencia alguna a cualquier funcionalidad, tamaño, estructuras tridimensionales u origen. Por lo tanto, a un fragmento o porción de una proteína Cry2A de la invención se le sigue aludiendo en esta memoria como una "proteína".

De acuerdo con esta invención se han aislado y caracterizado secuencias de ácidos nucleicos, particularmente secuencias de ADN que codifican nuevas toxinas Cry de *Bt*. Los nuevos genes se designaron *cry2Ae*, *cry2Af*, *cry2Ag* y sus proteínas codificadas Cry2Ae, Cry2Af y Cry2Ag.

De acuerdo en esta invención, "proteína Cry2Ae" se refiere a cualquier proteína comparativa que comprende el fragmento más pequeño de la secuencia de aminoácidos de SEQ ID N° 2 que conserva la actividad insecticida (al que se alude en esta memoria como "fragmento tóxico más pequeño"), particularmente cualquier proteína que comprenda la secuencia de aminoácidos desde el aminoácido en la posición 1 al aminoácido en la posición 625, particularmente al aminoácido en la posición 632 en SEQ ID N° 2. Esto incluye híbridos o proteínas quiméricas que comprenden el fragmento de proteína tóxica más pequeño, así como proteínas que contienen al menos uno de los tres dominios de la proteína de SEQ ID N° 2. También se incluyen en esta definición variantes de la secuencia de aminoácidos en SEQ ID N° 2 tales como proteínas con una identidad de la secuencia de al menos 92%, particularmente de al menos 93%, 95%, 96%, 97%, 98% o 99% al nivel de la secuencia de aminoácidos, según se determina utilizando alineamientos apareados utilizando el programa GAP del paquete Wisconsin de GCG (Madison, Wisconsin, EE.UU., versión 10.0; utilizar los valores por defecto GCG en el programa GAP; para las comparaciones de secuencias de aminoácidos, utilizar la matriz de registro blosum62), preferiblemente proteínas con algunos, preferiblemente 5-10, particularmente menos de 5 aminoácidos añadidos, reemplazados o suprimidos, sin modificar significativamente, preferiblemente sin modificar la actividad insecticida de la proteína, p. ej. la proteína Cry2Ae de SEQ ID N° 8.

La expresión "ADN/proteína que comprende la secuencia X", tal como se utiliza en esta memoria, se refiere a un ADN o proteína que incluye o contiene al menos la secuencia X, de modo que otras secuencias de nucleótidos o aminoácidos pueden incluirse en el extremo 5' (o N-terminal) y/o 3' (o C-terminal), p. ej. (la secuencia de nucleótidos de) una proteína marcadora seleccionable según se describe en el documento EP 0 193 259, (la secuencia de nucleótidos de) un péptido de tránsito y/o una secuencia conductora 5' ó 3'.

El "fragmento tóxico más pequeño" de una proteína Cry de la invención, tal como se utiliza en esta memoria, es el fragmento o porción más pequeño de una proteína Cry que conserva actividad insecticida que puede obtenerse mediante digestión enzimática, preferiblemente con tripsina o quimotripsina, de la proteína Cry de longitud completa, o el fragmento o porción más pequeño de una proteína Cry que conserva la actividad insecticida que puede obtenerse realizando deleciones de nucleótidos en el ADN que codifica una proteína Cry. Los extremos N- y C-terminales de la secuencia de aminoácidos del fragmento tóxico más pequeño se determinan convenientemente mediante determinación de la secuencia de aminoácidos de los fragmentos anteriores mediante técnicas rutinariamente disponibles en la técnica. Para los fragmentos de proteína Cry2A que conservan actividad insecticida de esta invención, típicamente se pueden realizar deleciones N-terminales, mientras que poco se puede suprimir en su extremo C-terminal. Para las proteínas Cry2Af de la invención, se espera que se puedan realizar

delecciones hasta la posición 625 del aminoácido en el extremo C (es decir, el aminoácido C-terminal sería el aminoácido en la posición 625), al tiempo que se conserva la actividad insecticida, para la proteína Cry2Ag se espera que se puedan realizar delecciones hasta la posición 620 del aminoácido en el extremo C (es decir, el aminoácido C-terminal sería el aminoácido en la posición 620) al tiempo que se conserve la actividad insecticida de la proteína. Es de esperar que las delecciones N-terminales hasta en torno a la posición 50 del aminoácido, preferiblemente delecciones N-terminales hasta la posición 50 del aminoácido (es decir, el aminoácido N-terminal estaría en la posición 50 de las secuencias mostradas en el listado de secuencias) en la secuencia de aminoácidos de las proteínas Cry2A de esta invención conserven la mayor parte de su actividad insecticida contra insectos lepidópteros.

De acuerdo con esta invención, "proteína Cry2Af" se refiere a cualquier proteína que comprende el fragmento tóxico más pequeño de la secuencia de aminoácidos de SEQ ID N° 4, particularmente cualquier proteína que comprenda la secuencia de aminoácidos desde el aminoácido en la posición 1 al aminoácido en la posición 625, particularmente al aminoácido en la posición 632 en SEQ ID N° 4. Esto incluye híbridos o proteínas quiméricas que comprenden el fragmento de proteína tóxico más pequeño, así como proteínas que contienen al menos uno de los tres dominios de la proteína de SEQ ID N° 4. También se incluyen en esta definición variantes de la secuencia de aminoácidos en SEQ ID N° 4 tales como proteínas con una identidad de la secuencia de al menos el 95%, particularmente de al menos el 97%, al menos el 98% o al menos el 99% al nivel de la secuencia de aminoácidos, según se determina utilizando alineamientos apareados utilizando el programa GAP del paquete Wisconsin de GCG (Madison, Wisconsin, EE.UU., versión 10.0; utilizar los valores por defecto GCG en el programa GAP; para las comparaciones de secuencias de aminoácidos, utilizar la matriz de registro blosum62), preferiblemente proteínas con algunos, preferiblemente 5-10, particularmente menos de 5 aminoácidos añadidos, reemplazados o suprimidos, sin modificar significativamente, preferiblemente sin modificar la actividad insecticida de la proteína.

De acuerdo con esta invención, "proteína Cry2Ag" se refiere a cualquier proteína comparativa que comprende el fragmento tóxico más pequeño de la secuencia de aminoácidos de SEQ ID N° 6, particularmente cualquier proteína que comprenda la secuencia de aminoácidos desde el aminoácido en la posición 1 al aminoácido en la posición 620, particularmente al aminoácido en la posición 627 en SEQ ID N° 6. Esto incluye híbridos o proteínas quiméricas que comprenden el fragmento de proteína tóxico más pequeño, así como proteínas que contienen al menos uno de los tres dominios de la proteína de SEQ ID N° 6. También se incluyen en esta definición variantes de la secuencia de aminoácidos en SEQ ID N° 6 tales como proteínas con una identidad de la secuencia de al menos el 80%, particularmente de al menos el 85%, el 90%, el 95%, el 96%, el 97%, el 98% o al menos el 99% al nivel de la secuencia de aminoácidos, según se determina según se determina utilizando alineamientos apareados utilizando el programa GAP del paquete Wisconsin de GCG (Madison, Wisconsin, EE.UU., versión 10.0; utilizar los valores por defecto GCG en el programa GAP; para las comparaciones de secuencias de aminoácidos, utilizar la matriz de registro blosum62), preferiblemente proteínas con algunos, preferiblemente 5-10, particularmente menos de 5 aminoácidos añadidos, reemplazados o suprimidos, sin modificar significativamente, preferiblemente sin modificar la actividad insecticida de la proteína.

Tal como se utiliza en esta memoria, las expresiones "ADN de *cry2Ae*", "ADN de *cry2Af*" o "ADN de *cry2Ag*" se refieren a cualquier secuencia de ADN que codifica la proteína Cry2Ae, Cry2Af o Cry2Ag, respectivamente, según se define arriba. Esto incluye secuencias de ADN que se producen de forma natural, artificiales o sintéticas que codifican las proteínas de SEQ ID N°s 2, 4 ó 6 o sus fragmentos o variantes insecticidas según se define arriba. También se incluyen en esta memoria secuencias de ADN que codifican proteínas insecticidas que son lo suficientemente similares a las regiones codificantes de las secuencias de ADN genómicas depositadas o a las secuencias proporcionadas en el listado de secuencias de modo que éstas pueden (es decir, tienen la capacidad de) hibridarse a estas secuencias de ADN bajo condiciones de hibridación rigurosas. Condiciones de hibridación rigurosas, tal como se utiliza en esta memoria, se refiere particularmente a las siguientes condiciones: inmovilizar las secuencias de ADN genómico relevantes sobre un filtro y pre-hibridar los filtros durante 1 a 2 horas en formamida al 50%, SSPE al 5%, reactivo 2x Denhardt y SDS al 0,1% a 42°C, o durante 1 a 2 horas en 6x SSC, reactivo 2x Denhardt y SDS al 0,1% a 68°C. La sonda marcada con dig o radiomarcada desnaturalizada se añade luego directamente al fluido de pre-hibridación y se lleva a cabo la incubación durante 16 a 24 horas a la temperatura apropiada arriba mencionada. Después de la incubación, los filtros se lavan luego durante 30 minutos a la temperatura ambiente en 2x SSC, SDS al 0,1%, seguido de 2 lavados de 30 minutos, cada uno a 68°C en 0,5x SSC y SDS al 0,1%. Se establece una autorradiografía exponiendo los filtros durante 24 a 48 horas a una película de rayos X (Kodak XAR-2 o equivalente) a -70°C con una pantalla intensificadora. Naturalmente, en este proceso se pueden utilizar condiciones y parámetros equivalentes al tiempo que se siguen conservando las condiciones de hibridación rigurosas deseadas. Variantes preferidas del ADN de *cry2Af* de esta invención son un ADN que codifica

las variantes de la proteína Cry2Af insecticida arriba descritas, o una secuencia de ADN que codifica una proteína insecticida con una identidad de al menos el 95%, preferiblemente al menos el 96% o 97%, más preferiblemente al menos el 98% o al menos el 99% con la secuencia codificante de SEQ ID N° 3. Particularmente, secuencias de ADN de este tipo también se hibridan bajo condiciones de hibridación rigurosas a la secuencia codificante de *cry2Af* depositada en la BCCM-LMBP bajo el número de acceso LMBP 4247 o a la secuencia codificante de SEQ ID N° 3. Las identidades de secuencias aludidas anteriormente se calculan utilizando el programa GAP del paquete Wisconsin de GCG (Madison, Wisconsin, EE.UU.), versión 10.0 (se utilizan los valores por defecto GCG, para las comparaciones de estas secuencias de ADN se utiliza la matriz de registro "nwsgapdna"), las condiciones de hibridación rigurosas son como se definen arriba.

"Actividad insecticida" de una proteína, tal como se utiliza en esta memoria, significa la capacidad de una proteína de exterminar insectos cuando dicha proteína se proporciona a insectos, preferiblemente mediante la expresión en un hospedante recombinante tal como una planta. "Cantidades represoras de insectos" de una proteína, tal como se utiliza en esta memoria, se refiere a una cantidad de proteína que es suficiente para limitar el daño a una planta por parte de insectos que se alimentan de dicha planta a niveles comercialmente aceptables, p. ej. exterminando los insectos o inhibiendo el desarrollo, fertilidad o crecimiento de los insectos de manera que esto proporcionen un menor daño a una planta y la cosecha no se vea significativamente afectada de manera adversa.

De acuerdo con esta invención, insectos susceptibles a las nuevas proteínas Cry de la invención se ponen en contacto con esta proteína en cantidades represoras de los insectos, preferiblemente cantidades insecticidas. Insectos diana preferidos para las proteínas de esta invención son plagas de insectos de perjuicio económico de plantas de maíz, algodón, arroz y soja, particularmente en los países de Norteamérica y Sudamérica. Insectos diana particularmente preferidos para las proteínas Cry2A de esta invención son *Heliothis* spp., *Helicoverpa* spp., *Spodoptera* spp., *Sesamia* spp., *Anticarsia* spp., *Ostrinia* spp., *Chilo* spp., *Sesamia* spp., *Marasmia* spp., *Scirpophaga* spp y *Cnaphalocrocis* spp, preferiblemente, más preferiblemente *Heliothis virescens*, *Helicoverpa zea*, *Helicoverpa armigera*, *Anticarsia gemmatalis* y *Ostrinia nubilalis*.

Las expresiones "proteína Cry2A", "proteína Cry2A de esta invención", "proteína Cry" o "proteína Cry", tal como se utilizan en esta memoria se refieren a una cualquiera de las nuevas proteínas aisladas de acuerdo con esta invención e identificadas y definidas en esta memoria como proteína Cry2Ae, Cry2Af o Cry2Ag. Una proteína Cry, tal como se utiliza en esta memoria, puede ser una proteína en su tamaño de longitud completa, también denominada una protoxina, o puede estar en una forma truncada en tanto que se conserve la actividad insecticida, o puede ser una combinación de diferentes proteínas en una proteína híbrida o de fusión. Una "protoxina Cry" se refiere a la proteína cristal de longitud completa según es codificada por la secuencia de ADN de *Bt* que se produce de forma natural, una "toxina Cry" se refiere a un fragmento insecticida de la misma, particularmente el fragmento tóxico más pequeño de la misma, típicamente en el intervalo de pesos moleculares de aproximadamente 50-65 kD, particularmente de aproximadamente 60 kD, según se determina mediante electroforesis en SDS-PAGE. Un "gen *cry*", "gen *cry2A*", "ADN de *cry*" o ADN de *cry2A*" tal como se utiliza en esta memoria, es una secuencia de ADN que codifica una proteína Cry de acuerdo con esta descripción, que se refiere a cualquiera de las secuencias de ADN *cry2Ae*, *cry2Af* o *cry2Ag* definidas anteriormente.

La secuencia de ácidos nucleicos, particularmente la secuencia de ADN, que codifica las proteínas Cry descritas en esta memoria se puede aislar de una manera convencional a partir de las cepas de *E. coli* recombinantes depositadas de acuerdo con el Tratado de Budapest el 6 de octubre del 2000 en la Vakgroep voor Moleculaire Biologie-Plasmidencollectie, Universiteit Gent, K.L. Ledeganckstraat 35, B-9000 Gent, Bélgica (abreviado aquí en lo que sigue como "BCCM-LMBP") bajo los siguientes números de acceso; BCCM-LMBP 4247 para la cepa XL1Blue:pUC1099E/*cry2clone1*, que codifica la proteína Cry2Af; BCCM-LMBP 4248 para la cepa XL1Blue:pUC1099E/*cry2clone7*, que codifica la proteína Cry2Ae; y BCCM-LMBP 4249 para la cepa XL1Blue:pUC2761A/*cry2clone141*, que codifica la proteína Cry2Ag. Las secuencias de ADN que codifican las proteínas Cry de la invención se pueden aislar a partir de estas cepas depositadas utilizando técnicas rutinarias y se pueden insertar en vectores de expresión para producir elevadas cantidades de proteínas Cry. Las proteínas Cry se pueden utilizar para preparar anticuerpos monoclonales o policlonales específicos de una manera convencional (Höfte et al., 1988).

También, las secuencias de ADN para uso en esta invención se pueden preparar de modo sintético. De hecho, debido a la degeneración del código genético, algunos codones de aminoácidos pueden ser reemplazados por otros sin modificar la secuencia de aminoácidos de la proteína. Además de ello, algunos aminoácidos pueden ser sustituidos por otros aminoácidos equivalentes sin modificar significativamente, preferiblemente sin modificar la actividad insecticida de la proteína. También cambios en la secuencia de aminoácidos o en la composición en

regiones de la molécula, diferentes de los responsables de la unión o de la formación de poros, son menos probables de provocar una diferencia en la actividad insecticida de la proteína. Equivalentes de las secuencias de ADN de la invención incluyen secuencias de ADN que se hibridan a la secuencia de ADN de las proteínas Cry de SEQ ID N° 3 bajo condiciones de hibridación rigurosas y que codifican una proteína con las mismas características insecticidas que la proteína de esta invención, o secuencias de ADN que tienen una diferente utilización del codón en comparación con los genes *cry2A* nativos de esta invención, pero que codifican una proteína con la misma actividad insecticida y con esencialmente la misma, preferiblemente la misma secuencia de aminoácidos. Ejemplos comparativos de secuencias de ADN optimizadas en los codones para la proteína Cry2Ae de esta invención se encuentran en las SEQ ID N°s 7 y 9. Estas secuencias de ADN se optimizaron adaptando la utilización de codones al más preferido en genes de plantas, particularmente a genes nativos para el género o especie de la planta de interés (Bennetzen y Hall, 1982; Itakura et al., 1977) utilizando tablas de utilización de codones disponibles (la SEQ ID N° 7 estaba más adaptada hacia la expresión en algodón, la SEQ ID N° 9 más hacia el maíz) y también para eliminar tramos de nucleótidos AT o GC más largos que 5 ó 6, preferiblemente más largos que 5 nucleótidos, y también para insertar sitios de restricción adecuados.

También, el extremo N de una proteína Cry se puede modificar de modo que tenga un contexto de iniciación de la traducción óptimo, añadiendo o suprimiendo con ello uno o más aminoácidos en el extremo N-terminal de la proteína. En la mayoría de los casos, se prefiere que las proteínas de la invención sean expresadas en células vegetales, comenzando con un dipéptido Met-Asp o Met-Ala para la iniciación óptima de la traducción, requiriendo la inserción en el ADN de *cry2A* de un codón que codifique un aminoácido Asp o Ala más abajo del codón de iniciación en forma de un nuevo segundo codón.

Naturalmente, puede construirse cualquier secuencia de ADN que difiera en su utilización del codón, pero que codifique la misma proteína o una proteína similar con esencialmente la misma actividad insecticida, dependiendo del fin particular. Se ha descrito en sistemas de expresión procarióticos y eucarióticos que se desea el cambio de la utilización del codón por el de la célula hospedante para la expresión de genes en hospedantes extraños (Bennetzen y Hall, 1982, Itakura et al., 1977). Además de ello, se sabe que genes de la proteína cristal de *Bt* no tienen desviación alguna hacia los codones eucarióticos y que son muy ricos en AT (Adang et al., 1985, Schnepf et al., 1985). En la bibliografía están disponibles tablas de utilización de codones (Wada et al., 1990; Murray et al., 1989) y en las bases de datos de secuencias de ADN principales (p. ej. EMBL en Heidelberg, Alemania). Por consiguiente, se pueden construir secuencias de ADN sintético de modo que se produzcan las mismas o esencialmente las mismas proteínas. Es evidente que se pueden producir varias secuencias de ADN una vez que se conoce la secuencia de aminoácidos de las proteínas Cry de esta invención. Estas otras secuencias de ADN incluyen secuencias de ADN sintéticas o semi-sintéticas que han sido modificadas con el fin de inactivar determinados sitios en el gen, p. ej. inactivando selectivamente determinados elementos reguladores o de procesamiento presentes en la secuencia nativa según se describe en las publicaciones PCT WO 91/16432 y WO 93/09218, o adaptando la utilización general de codones al de un organismo hospedante más relacionado, preferiblemente al del organismo hospedante en el que se desee la expresión. En la bibliografía de patentes y científica se pueden encontrar varias técnicas para modificar la utilización de codones al preferido por las células hospedantes. El método exacto de la modificación de la utilización de codones no es crítico para esta invención en tanto que la mayor parte o la totalidad de las secuencias reguladoras cripticas o elementos de procesamiento hayan sido reemplazados por otras secuencias. Ejemplos de secuencias de ADN utilizadas para la expresión en plantas se muestran en las SEQ ID N°s 7 y 9 adjuntas.

Pequeñas modificaciones a una secuencia de ADN tal como la descrita anteriormente se pueden realizar rutinariamente, es decir, mediante mutagénesis mediada por PCR (Ho et al., 1989, White et al., 1989). Modificaciones más profundas a una secuencia de ADN se pueden realizar rutinariamente mediante la síntesis de ADN de novo de una región codificante deseada utilizando técnicas disponibles.

Con la expresión "esencialmente la misma", cuando se alude a la secuencia de aminoácidos de una proteína Cry, se quiere dar a entender que incluye una secuencia de aminoácidos que no difiere en más del 5%, preferiblemente en no más del 2% de la secuencia de aminoácidos de la proteína con la que se compara; y cuando se alude a toxicidad de una proteína Cry, se quiere dar a entender que incluye una proteína, cuyo valor LC₅₀ obtenido bajo las mismas condiciones de bioensayo difiere en no más de un 10%, preferiblemente en no más de un 5% del valor LC₅₀ obtenido para la proteína con la que se compara.

El término "dominio" de una toxina Cry, tal como se utiliza en esta memoria, significa cualquier o cualesquiera partes o dominios de la toxina con una estructura específica que puede ser transferida a otra proteína (Cry) para

proporcionar una nueva proteína híbrida con al menos una característica funcional (p. ej. las características de unión y/o toxicidad) de la toxina Cry de la invención (Ge et al., 1991). Dichas partes pueden formar una característica esencial de la proteína híbrida de *Bt* con las características de unión y/o toxicidad de la proteína Cry de esta invención. Una proteína híbrida de este tipo puede tener un intervalo de hospedantes ampliado, una toxicidad mejorada y/o puede utilizarse en una estrategia para prevenir el desarrollo de resistencia a los insectos (publicación de patente europea ("EP") 408 403; Visser et al., 1993).

Las secuencias de ADN de *cry* de la invención, preparadas a partir de ADN total, se pueden ligar en vectores de expresión adecuados y se pueden transformar en *E. coli*, y los clones pueden luego ser rastreados mediante métodos convencionales de inmuno-sondeo de colonias (French et al., 1986) para la expresión de la toxina con anticuerpos monoclonales o policlonales dirigidos contra las proteínas Cry.

También, el ADN de *cry* de la invención se puede ligar en vectores lanzadera de *Bt* adecuados (Lereclus et al., 1992) y se pueden transformar en un mutante cristal-minus de *Bt*. Los clones se pueden luego rastrear en cuanto a la producción de cristales (detectados por microscopía) o proteínas cristal (detectadas mediante SDS-PAGE), o se puede someter a ensayo en cuanto a su actividad insecticida en comparación con la cepa cristal-minus testigo.

Los genes que codifican las proteínas Cry de esta invención se pueden secuenciar de una manera convencional (Maxam y Gilbert, 1980; Sanger, 1977), para obtener la secuencia de ADN. Comparaciones de la secuencia indicaban que los genes son diferentes de los genes previamente descritos que codifican protoxinas y toxinas con actividad contra lepidópteros (véase, p. ej., Höfte y Whiteley, 1989; Crickmore, et al., 1998; y la actualización de 16 de octubre de 2000 de la página web de la nomenclatura de *Bt* correspondiente a la publicación de Crickmore et al. (1998), encontrada en: http://epunix.biols.susx.ac.uk/Home/Neil_Crickmore/Bt/index.html). También, las proteínas Cry2A de la invención son nuevas frente a cualquiera de las secuencias de proteínas cristal de *Bacillus thuringiensis* en la actualización del 13 de diciembre de 2001 de esta página web de la nomenclatura de *Bt*.

Una parte insecticidamente eficaz de las secuencias de ADN, que codifican una parte insecticidamente eficaz de las formas protoxina de proteína Cry recientemente identificadas, se puede realizar de una manera convencional después del análisis de la secuencia del gen. En fragmentos de este tipo, se prefiere que al menos la secuencia homóloga al bloque 5 de la secuencia conservada de proteínas cristal de *Bt* (Höfte y Whiteley, 1989; Schnepf et al., 1998) esté incluida en una proteína de este tipo, preferiblemente hasta dos aminoácidos después de esta región homóloga. Para proteínas Cry2Af, esta región homóloga termina en la posición de aminoácido 625 en SEQ ID N° 4, para la proteína Cry2Ag comparativa en la posición 620 en la SEQ ID N° 6. La secuencia de aminoácidos de las proteínas Cry se puede determinar a partir de la secuencia de ADN de las secuencias de ADN aisladas. Por "una parte (o porción o fragmento) insecticidamente eficaz" de secuencias de ADN que codifican la proteína Cry, a las que se alude también en esta memoria como "gen truncado" o "ADN truncado" se quiere dar a entender una secuencia de ADN que codifica un polipéptido que tiene menos aminoácidos que la forma protoxina de la proteína Cry, pero que es insecticida.

Con el fin de expresar la totalidad o una parte insecticidamente eficaz de la secuencia de ADN que codifica una proteína Cry de esta invención en *E. coli*, en otras cepas de *Bt* y en plantas, se pueden introducir sitios de restricción adecuados que flanquean la secuencia de ADN. Esto se puede realizar mediante mutagénesis dirigida al sitio utilizando procesos bien conocidos (Stanssens et al., 1989; White et al., 1989). Con el fin de obtener una expresión mejorada en plantas, la utilización de codones del gen *cry* o parte del gen *cry* insecticidamente eficaz de esta invención se puede modificar para formar un gen o parte de un gen equivalente, modificada o artificial de acuerdo con las publicaciones PCT WO 91/16432 y WO 93/09218; EP 0 358 962 y EP 0 359 472, o los genes o partes de genes de *Bt* se pueden insertar en el genoma del plástido, mitocondrial o cloroplasto y expresar allí utilizando un promotor adecuado (p. ej., Mc Bride et al., 1995; patente de EE.UU. 5.693.507). Para obtener una expresión potenciada en plantas monocotiledóneas tales como maíz, también se puede añadir al gen quimérico un intrón, preferiblemente un intrón monocotiledóneo, y la secuencia de ADN del gen *cry* o de su parte insecticida se puede modificar adicionalmente de una manera neutra en cuanto a la traducción, para modificar, posiblemente inhibir las secuencias de ADN presentes en la parte del gen por medio de inserción de intrones dirigida al sitio y/o introduciendo cambios al uso de codones, p. ej. adaptando la utilización de codones al más preferido por las plantas, preferiblemente al género de plantas relevante específico (Murray et al., 1989) sin modificar significativamente, preferiblemente sin modificar la secuencia de aminoácidos codificada.

De acuerdo con una realización de esta invención, se prefiere que las proteínas estén dirigidas a organelas intracelulares tales como plástidos, preferiblemente cloroplastos, mitocondrias, o sean secretadas a partir de la

célula, optimizando potencialmente la estabilidad y/o expresión de proteínas. Para este fin, los genes quiméricos de la invención comprenden una región codificante, que codifica un péptido señal o diana, enlazada a la región codificante de la proteína Cry de la invención. Péptidos particularmente preferidos a incluir en las proteínas de esta invención son los péptidos tránsito para la fijación como objetivo de cloroplastos u otros plástidos, especialmente regiones de péptidos de tránsito duplicadas procedentes de genes de plantas cuyo producto génico está fijado como objetivo a los plástidos, el péptido de tránsito optimizado de Capellades et al. (patente de EE.UU. 5.635.618), el péptido de tránsito de ferredoxina-NADP⁺ oxidoreductasa de espinacas (Oelmuller et al., 1993), el péptido de tránsito descrito en Wong et al. (1992) y los péptidos diana en la solicitud de patente PCT publicada WO 00/26371. También se prefieren péptidos que señalizan la secreción de una proteína enlazada a este tipo de péptidos fuera de la célula tal como la señal de secreción del inhibidor de proteinasa II de patatas (Keil et al., 1986), la señal de secreción de gen alfa-amilasa 3 del arroz (Sutliff et al., 1991) y la señal de secreción de la proteína del tabaco PR1 (Cornelissen et al., 1986).

Péptidos señal particularmente útiles de acuerdo con la invención, incluyen el péptido de tránsito de cloroplastos (p. ej. Van Den Broeck et al. (1985) o el péptido de tránsito de cloroplastos de la patente de EE.UU. 5.510.471 y la patente de EE.UU. 5.635.618 que provoca el transporte de la proteína a los cloroplastos, un péptido señal secretor o un péptido que fija como objetivo la proteína a otros plástidos, mitocondrias, el ER (retículo endoplasmático) u otras organelas. Secuencias señal para fijar como objetivo organelas intracelulares o para la secreción fuera de la célula vegetal o a la pared celular se encuentran en proteínas fijadas naturalmente como objetivo o secretadas, preferiblemente las descritas por Klösgen et al. (1989), Klösgen y Weil (1991), Neuhaus y Rogers (1998), Bih et al. (1999), Morris et al. (1999), Hesse et al. (1989), Tavladoraki et al. (1998), Terashima et al. (1999), Park et al. (1997), Shcherban et al. (1995), todas las cuales se incorporan en esta memoria como referencia, particularmente las secuencias de péptidos señal procedentes de proteínas fijadas como objetivo o secretadas de maíz, algodón, arroz o soja.

Además de ello, las propiedades de unión de las proteínas Cry de la invención se pueden evaluar utilizando métodos conocidos en la técnica (p. ej. Van Rie et al., 1990), para determinar si las proteínas Cry de la invención se unen a sitios en el tracto digestivo de los insectos que no son reconocidos (o por los que se compite) por otras proteínas Cry conocidas u otras proteínas de *Bt*. Toxinas *Bt* con diferentes sitios de unión para los que existe una unión no competitiva en insectos susceptibles relevantes son muy valiosos para reemplazar toxinas de *Bt* conocidas para las cuales los insectos pudieran haber desarrollado resistencia, o para uso en combinación con toxinas de *Bt* con un diferente modo de acción para prevenir o demorar el desarrollo de la resistencia de los insectos frente a toxinas de *Bt*, particularmente cuando se expresan en una planta. Debido a las características de las toxinas de *Bt* recientemente aisladas, éstas son extremadamente útiles para transformar plantas, p. ej. monocotiledóneas tales como plantas de maíz o arroz, y dicotiledóneas tales como plantas de algodón, soja y especies de *Brassica*, para proteger a estas plantas frente a un daño por insectos. Se ha descrito que en *Helicoverpa zea*, la proteína Cry2Aa no comparte sitios de unión con la proteína Cry1Ac (English et al., 1994). De manera similar, se espera que las propiedades de unión de las proteínas Cry2A de la presente invención sean diferentes en comparación con las de toxinas Cry1 o Cry9 actualmente utilizadas en plantas transgénicas en las plagas de insectos relevantes. Propiedades de unión diferentes de este tipo se pueden medir mediante ensayos de unión rutinarios según se describe antes. Especialmente para fines de gestión de la resistencia a insectos para una plaga de insectos específica, se prefiere combinar una proteína Cry2A de esta invención con otra proteína testigo de insectos, particularmente una proteína cristal de Bt, que no reconoce al menos un sitio de unión reconocido por una proteína Cry2A de este tipo. Proteínas testigo de insectos preferidas a combinar con las proteínas Cry2A de esta invención, particularmente para la expresión simultánea en plantas, preferiblemente plantas de algodón, incluyen la proteína Cry1F o híbridos derivados de una proteína Cry1F (p. ej. las proteínas Cry1A-Cry1F híbridas descritas en las patentes de EE.UU. 6.326.169; 6.281.016; 6.218.188 o fragmentos tóxicos de las mismas), las proteínas de tipo Cry1A o fragmentos tóxicos de las mismas, preferiblemente la proteína Cry1Ac o híbridos derivados de la proteína Cry1Ac (p. ej. la proteína híbrida Cry1Ab-Cry1Ac descrita en la patente de EE.UU. 5.880.275), la proteína VIP3Aa o un fragmento tóxico de la misma según se describe en Estruch et al., 1996 y la patente de EE.UU. 6.291.156, proteínas insecticidas de cepas de especies de *Xenorhabdus*, *Serratia* o *Photorhabdus* (p. ej. Waterfield et al., 2001; French-Constant y Bowen, 2000). En una realización, una co-expresión de este tipo se obtiene fácilmente transformando una planta que ya expresa una proteína para la represión de insectos con una Cry2A de esta invención, o cruzando plantas transformadas con la proteína para la represión de insectos y plantas transformadas para la proteína Cry2A de esta invención. Métodos para obtener la expresión de diferentes proteínas insecticidas de Bt (o, similarmente, para otras proteínas para la represión de insectos) en la misma planta en un esfuerzo por minimizar o prevenir el desarrollo de resistencia a plantas resistentes a insectos transgénicas se describen en la patente EP 0 408 403.

Las proteínas Cry2A de esta invención también se pueden utilizar convenientemente para reprimir insectos en el caso en el que el insecto desarrolle resistencia contra proteínas para la represión de insectos tales como las proteínas Cry1 de Bt que actualmente ya se están comercializando en plantas transgénicas.

- 5 Preferiblemente, para fines de selección, pero también para aumentar las opciones de la represión de malas hierbas, las plantas transgénicas de la invención se transforman también con un ADN que codifica una proteína que confiere resistencia a un herbicida de amplio espectro, p. ej. herbicidas basados en glufosinato o glifosato.

10 La parte del gen *cry* insecticidamente eficaz o su equivalente, preferiblemente el gen quimérico *cry*, que codifica una porción insecticidamente eficaz de la protoxina Cry puede insertarse de forma estable de una manera convencional en el genoma nuclear de una sola célula vegetal, y la célula vegetal, así transformada, se puede utilizar de una manera convencional para producir una planta transformada que es resistente a insectos. A este respecto, un plásmido Ti desarmado, que contiene la parte del gen *cry* insecticidamente eficaz, en *Agrobacterium tumefaciens* se puede utilizar para transformar la célula vegetal y, después de ello, una planta transformada se puede regenerar a partir de la célula vegetal transformada utilizando los procesos descritos, por ejemplo, en los documentos EP 0 116 718, EP 0 270 822, publicación PCT WO 84/02913 y solicitud de patente europea publicada ("EP") 0 242 246 y en Gould et al. (1991). Vectores de plásmidos Ti preferidos contienen cada uno la parte del gen *cry* insecticidamente eficaz entre las secuencias límites, o al menos están situados hacia la izquierda de la secuencia del límite derecho, del ADN-T del plásmido Ti. Naturalmente, se pueden utilizar otros tipos de vectores para transformar la célula vegetal utilizando procesos tales como la transferencia de genes directa (según se describe, por ejemplo, en el documento EP 0 233 247). la transformación mediada por polen (según se describe, por ejemplo, en el documento EP 0 270 356, la publicación PCT WO 85/01856 y la patente de EE.UU. 4.684.611), la transformación mediada por virus de ARN de plantas (según se describe, por ejemplo en el documento EP 0 067 553 y en la patente de EE.UU. 4.407.956), la transformación mediada por liposomas (según se describe, por ejemplo, en la patente de EE.UU. 4.536.475 y otros métodos tales como los métodos recientemente descritos para transformar determinadas líneas de maíz (p. ej. patente de EE.UU. 6.140.553; Fromm et al., 1990; Gordon-Kamm et al., 1990) y de arroz (Shimamoto et al., 1989; Datta et al., 1990) y el método para transformar monocotiledóneas en general (publicación PCT WO 92/09696). Para la transformación de algodón se prefiere especialmente el método descrito en la publicación de patente PCT WO 00/71733. Para la transformación de soja, se hace referencia a métodos conocidos en la técnica, p. ej. Hinchee et al. (1988) y Christou et al. (1990) o el método del documento WO 00/42207.

También, además de la transformación del genoma nuclear, en la invención se incluye la transformación del genoma del plástido, preferiblemente el genoma de cloroplastos. Kota et al. (1999) han descrito un método para sobre-expresar una proteína Cry2Aa en cloroplastos de tabaco.

40 La planta transformada resultante se puede utilizar en un esquema de cultivo de plantas convencional para producir plantas más transformadas con las mismas características o para introducir la parte del gen *cry* herbicidamente eficaz en otras variedades de la misma o de especies vegetales relacionadas. Semillas, que se obtienen a partir de las plantas transformadas, contienen la parte del gen *cry* insecticidamente eficaz en forma de un inserto genómico estable. Células de la planta transformada se pueden cultivar de una manera convencional para producir la porción insecticidamente eficaz de la protoxina Cry, preferiblemente la toxina Cry, que se puede recuperar para uso en composiciones insecticidas convencionales contra lepidópteros (patente de EE.UU. 5.254.799).

45 La parte del gen *cry* insecticidamente eficaz, preferiblemente el gen *cry* truncado, se inserta en el genoma de una célula vegetal de modo que el gen insertado se encuentra situado más abajo (es decir, 3') y, bajo el control de un promotor que puede dirigir la expresión de la parte de gen en la célula vegetal. Esto se consigue preferiblemente insertando el gen quimérico *cry* en el genoma de la célula vegetal, particularmente en el genoma nuclear o del plástido (p. ej. cloroplasto). Promotores preferidos incluyen: los promotores 35S fuertes constitutivos (los "promotores 35S") del virus del mosaico de la coliflor (CaMV) de aislados CM 1841 (Gardner et al., 1981), CabbB-S (Franck et al., 1980) y CabbB-JI (Hull y Howell, 1987); el promotor 35S descrito por Odell et al. (1985), promotores de la familia ubiquitina (p. ej., el promotor de la ubiquitina del maíz de Christensen et al., 1992, véase también Cornejo et al., 1993), el promotor *gos2* (de Pater et al., 1992), el promotor *emu* (Last et al., 1990), promotores actina de *Arabidopsis* tales como el promotor descrito por An et al. (1996), promotores actina del arroz tales como el promotor descrito por Zhang et al., (1991); promotores del virus del mosaico de las nervaduras de la mandioca (documento WO 97/48819, Verdaguer et al. (1998)), la serie pPLEX de promotores procedentes del virus de la atrofia subterránea del trébol (documento WO 96/06932, particularmente el promotor S7), un promotor de

alcohol deshidrogenasa, p. ej. pAdh1S (números de acceso a GenBank X04049, X00581), y el promotor TR1' y el promotor TR2' (el "promotor TR1'" y "promotor TR2'", respectivamente) que impulsan la expresión de los genes 1' y 2', respectivamente, del ADN-T (Velten et al., 1984). Alternativamente, se puede utilizar un promotor que no sea constitutivo, sino más bien específico para uno o más tejidos u órganos de la planta (p. ej. hojas y/o raíces), con lo que la parte del gen *cry* insertada se expresa solamente en células del o de los tejidos u órganos específicos. Por ejemplo, la parte del gen *cry* insecticidamente eficaz podría ser selectivamente expresada en las hojas de una planta (p. ej. maíz, algodón) colocando la parte del gen insecticidamente eficaz bajo el control de un promotor inducible para la luz tal como el promotor del gen de la subunidad pequeña de ribulosa-1,5-bisfosfato carboxilasa de la planta propiamente dicha o de otra planta tal como el guisante, según se describe en la patente de EE.UU. 5.254.799. Otra alternativa consiste en utilizar un promotor, cuya expresión sea inducible, preferiblemente hiriendo tal como la alimentación de insectos, p. ej. el promotor MPI descrito por Cordera et al. (1994), o por factores químicos.

La parte del gen *cry* insecticidamente eficaz se inserta en el genoma de la planta de modo que la parte del gen insertada se encuentra más arriba (es decir 5') de las señales de regulación de la transcripción en el extremo 3' adecuadas (es decir, las señales de la formación de transcrito y de poliadenilación). Esto se consigue preferiblemente insertando el gen quimérico *cry* en el genoma de la célula vegetal. Señales de poliadenilación y de formación del transcrito preferidas incluyen las del gen nopláina sintasa (Depicker et al., 1982), el gen octopina sintasa (Gielen et al., 1984) y el gen 7 de ADN-T (Velten y Schell, 1985), que actúan como secuencias de ADN 3'-no traducidas en células vegetales transformadas.

La parte del gen *cry* insecticidamente eficaz puede insertarse opcionalmente en el genoma de la planta en forma de un gen híbrido (patente de EE.UU. 5.254.799; Vaeck et al., 1987) bajo el control del mismo promotor en forma de un gen marcador seleccionable o puntuable tal como el gen *neo* (documento EP 0 242 236) que codifica la resistencia a canamicina, de modo que la planta expresa una proteína de fusión que es fácilmente detectable.

La transformación de células vegetales también se puede utilizar para producir las proteínas de la invención en grandes cantidades en cultivos de células vegetales, p. ej. para producir una proteína Cry2A que luego se puede aplicar a cosechas después de una formulación adecuada. Cuando se hace referencia en esta memoria a una célula vegetal transgénica, esto se refiere a una célula vegetal (o también un protoplasto vegetal) como tal en cultivo de aislamiento o de tejidos, o a una célula vegetal (o protoplasto) contenida en una planta o en un órgano o tejido diferenciado, y ambas posibilidades se incluyen específicamente en esta memoria. Por lo tanto, una referencia a una célula vegetal en la descripción o las reivindicaciones no quiere referirse sólo a células aisladas en cultivo, sino que se refiere a cualquier célula vegetal donde quiera que esté localizada o en cualquier tipo de tejido u órgano de la planta en el que pueda estar presente.

La totalidad o parte del gen *cry* que codifica una proteína anti-lepidópteros también se puede utilizar para transformar otras bacterias tales como una *B. thuringiensis* que tiene actividad insecticida contra lepidópteros o coleópteros. Con ello, se puede producir una cepa de *Bt* transformada que sea útil para combatir un amplio espectro de plagas de insectos lepidópteros y coleópteros o para combatir plagas de insectos lepidópteros adicionales. La transformación de bacterias tales como bacterias del género *Pseudomonas*, *Agrobacterium*, *Bacillus* o *Escherichia* con la totalidad o parte del gen *cry* de esta invención, incorporada en un vehículo de clonación adecuado, puede ser llevada a cabo de una manera convencional, preferiblemente utilizando técnicas de electroporación convencionales según se describe en Mahillon et al., (1989) y en la publicación de patente PCT WO 90/06999.

Cepas de especies de *Bacillus* transformadas que contienen el gen *cry* de esta invención se pueden fermentar por métodos convencionales (Dulmage, 1981; Bernhard y Utz, 1993) para proporcionar altos rendimientos de células. Bajo condiciones apropiadas que son bien comprendidas (Dulmage, 1981), cada una de estas cepas esporula para producir proteínas cristal que contienen la protoxina Cry en altos rendimientos.

Una composición insecticida, particularmente anti-lepidópteros, de esta invención se puede formular de una manera convencional utilizando los microorganismos transformados con el gen *cry* o, preferiblemente, sus respectivas proteínas Cry o la protoxina, toxina o parte de protoxina Cry insecticidamente eficaz en calidad de un ingrediente activo, junto con soportes, diluyentes, emulsionantes y/o dispersantes adecuados (p. ej. según se describe por Bernhard y Utz, 1993). Esta composición insecticida se puede formular en forma de un polvo humectable, nódulos, gránulos o polvo espolvoreable o en forma de una formulación líquida con disolventes acuosos o no acuosos tales como espuma, gel, suspensión, concentrado, etc.

5 Un método para reprimir insectos, particularmente lepidópteros, de acuerdo con esta invención puede comprender aplicar (p. ej. atomizar) a un lugar (área) a proteger, una cantidad insecticida de las proteínas Cry o células hospedantes transformadas con el gen *cry* de esta invención. El lugar a proteger puede incluir, por ejemplo, el hábitat de las plagas de insectos o la vegetación en desarrollo o un área en la que se ha de desarrollar vegetación.

10 Esta invención se refiere, además, a un método para reprimir plagas de insectos lepidópteros de la soja, particularmente lepidópteros barrenadores del tallo del arroz, patrones de arroz, gusanos cortadores del arroz, gusanos cogolleros del arroz o enrolladores de la hoja del arroz, preferiblemente un insecto seleccionado del grupo que consiste en: *Chilo suppressalis*, *Chilo partellus*, *Scirpophaga incertulas*, *Sesamia inferens*, *Cnaphalocrocis medinalis*, *Marasmia patnalis*, *Marasmia exigua*, *Marasmia ruralis*, *Scirpophaga innotata*, método que comprende aplicar a una zona o planta a proteger una proteína Cry2A según se define en esta memoria (es decir, plantando una planta de arroz transformada con un gen *cry2A* de esta invención o atomizando una composición que contiene una proteína Cry2A de esta invención). La invención se refiere también al uso de las proteínas Cry2A de esta invención contra plagas de insectos lepidópteros del arroz para minimizar el daño a las plantas de arroz.

20 Esta invención se refiere, además, a un método para reprimir plagas de insectos lepidópteros del algodón, método que comprende aplicar a una zona o planta a proteger una proteína Cry2A según se define en esta memoria (es decir, plantando una planta de arroz transformada con un gen Cry2A de esta invención o atomizando una composición que contiene una proteína Cry2A de esta invención). La invención se refiere también al uso de las proteínas Cry2A de esta invención frente a plagas de insectos lepidópteros del arroz para minimizar el daño a las plantas de arroz.

25 La invención se refiere también a un método para reprimir plagas de insectos lepidópteros del arroz, particularmente lepidópteros barrenadores del tallo del arroz, patrones de arroz, gusanos cortadores del arroz, gusanos cogolleros del arroz o enrolladores de la hoja del arroz, preferiblemente un insecto seleccionado del grupo que consiste en: *Chilo suppressalis*, *Chilo partellus*, *Scirpophaga incertulas*, *Sesamia inferens*, *Cnaphalocrocis medinalis*, *Marasmia patnalis*, *Marasmia exigua*, *Marasmia ruralis*, *Scirpophaga innotata*, método que comprende aplicar a una zona o planta a proteger una proteína Cry2A según se define en esta memoria (es decir, plantando una planta de arroz transformada con un gen *cry2A* de esta invención o atomizando una composición que contiene una proteína Cry2A de esta invención). La invención se refiere también al uso de las proteínas Cry2A de esta invención contra plagas de insectos lepidópteros del arroz para minimizar el daño a las plantas de arroz.

35 Para obtener la protoxina o toxina Cry, células de los hospedantes recombinantes que expresan la proteína Cry pueden desarrollarse de una manera convencional en un medio de cultivo adecuado y luego pueden lisarse utilizando medios convencionales tales como degradación enzimática, o detergentes, o similar. La protoxina puede luego separarse y purificarse por técnicas convencionales tales como cromatografía, extracción, electroforesis o similares. La toxina se puede luego obtener mediante digestión con tripsina de la protoxina.

40 Estas y/u otras realizaciones de esta invención se reflejan en las reivindicaciones que forman parte de la descripción de la invención.

45 Los siguientes Ejemplos ilustran la invención y no se proporcionan para limitar la invención ni la protección buscada. El listado de secuencias al que se alude en los Ejemplos, las reivindicaciones y la parte descriptiva es como sigue:

Listado de secuencias:

50 SEQ ID N° 1 – secuencia de aminoácidos y de ADN de proteína Cry2Ae comparativa y ADN.
 SEQ ID N° 2 – secuencia de aminoácidos de proteína Cry2Ae comparativa.
 SEQ ID N° 3 – secuencia de aminoácidos y de ADN de proteína Cry2Af y ADN.
 SEQ ID N° 4 – secuencia de aminoácidos de proteína Cry2Af.
 SEQ ID N° 5 – secuencia de aminoácidos y de ADN de proteína Cry2Ag comparativa y ADN.
 55 SEQ ID N° 6 – secuencia de aminoácidos de proteína Cry2Ag comparativa.
 SEQ ID N° 7 – secuencia de ADN de *cry2Ae* artificial comparativa para la expresión en algodón.
 SEQ ID N° 8 – secuencia de aminoácidos de la proteína Cry2Ae comparativa codificada por el ADN de SEQ ID N° 7.

SEQ ID N° 9 – secuencia de ADN de *cry2Ae* artificial comparativa para la expresión en maíz.

A menos que se establezca de otro modo en los Ejemplos, todos los procesos para producir y manipular ADN recombinante se llevaron a cabo por los procesos convencionales descritos en Sambrook et al., Molecular Cloning – A Laboratory Manual, segunda ed., Cold Spring Harbor Laboratory Press, NY (1989) y en los Volúmenes 1 y 2 de Ausubel et al. (1994) Current Protocols in Molecular Biology, Current Protocols, EE.UU. Materiales y métodos convencionales para el trabajo de biología molecular de plantas se describen en Plant Molecular Biology Labfax (1993) por R.R.D. Croy, publicado conjuntamente por BIOS Scientific Publications Ltd (Reino Unido) y Blackwell Scientific Publications (Reino Unido). Procesos para la tecnología de PCR se pueden encontrar en “PCR protocols: a guide to methods and applications” editado por M.A. Innis, D.H. Gelfand, J.J. Sninsky y T.J. White (Academic Press, Inc., 1990).

EJEMPLOS

15 Ejemplo 1: Caracterización de las cepas

Las cepas BTS02761A Y BTS01099E se aislaron a partir de polvo de cereales recogido en las Filipinas (South Tagalog) y Bélgica (Deerlijk), respectivamente.

Cada una de las cepas se puede cultivar en medios estándar convencionales, preferiblemente medio T₃ (triptona 3 g/l, triptosa 2 g/l, extracto de levaduras 1,5 g/l, 5 mg de MnCl₂, Na₂HPO₄·2H₂O 0,05 M, NaH₂PO₄·H₂O 0,05 M, pH 6,8 y 1,5% de agar), preferiblemente a 28°C. Para el almacenamiento a largo plazo, se prefiere mezclar un volumen igual de una suspensión de esporas-cristal con un volumen igual de glicerol al 50% y almacenar esto a -70°C, o liofilizar una suspensión de esporas-cristal. Para la esporulación, se prefiere el crecimiento en medio T₃ durante 72 horas a 28°C, seguido de almacenamiento a 4°C. Las proteínas cristal producidas por parte de las cepas durante la esporulación se empaquetan en cristales.

30 Ejemplo 2: Actividad insecticida de las cepas BTS02761A y BTS01099E contra especies de insectos de lepidópteros seleccionadas

Se realizaron ensayos de toxicidad en larvas neonatas de *Helicoverpa zea*, *Helicoverpa armígera*, *Heliothis virescens*, *Ostrinia nubilalis*, *Spodoptera frugiperda* y *Sesamia nonagrioides* alimentadas con una dieta artificial cubierta con un extracto alcalino (pH 12) no diluido de mezclas de esporas y cristales procedentes de BTS01099E o BTS02761A.

La dieta artificial (Vanderzant, 1962) se dispensó en pocillos de placas Costar de 48 pocillos. 25 microlitros del extracto sobre la superficie de la dieta se secaron en un flujo de aire laminar. En cada uno de los pocillos se colocó una larva y se utilizaron 18 larvas por cada muestra. Se recontaron al séptimo día las larvas muertas y vivas. El porcentaje de larvas muertas se muestra en la Tabla I que figura más abajo.

40 Mezclas de esporas/cristales procedentes de cada una de las cepas BTS02761A y BTS01099E se testaron en bioensayos y proporcionaron los siguientes resultados:

Tabla I:

Cepa	Mortalidad (%)				
	H _z	H _v	S _f	O _n	S _n
BTS02761A	17*	94	5	88	77
BTS01099E	70	100	NT	90	NT

45 *: larvas supervivientes ligeramente afectadas en su crecimiento

Testigos negativos (dieta estándar): H_z: 6% M, H_v: 17% M, S_f: 0% M.

H_z: *Helicoverpa zea*; H_v: *Heliothis virescens*; S_f: *Spodoptera frugiperda*; O_n: *Ostrinia nubilalis*; S_n: *Sesamia nonagrioides* (NT no testados).

50 Ejemplo 3: Identificación y caracterización de nuevos genes *cry2A* procedentes de las cepas BTS01099E y BTS02761A de Bt

Utilizando cebadores apropiados, se amplificó una porción del o de los genes *cry2A* de las cepas BTS02761A y BTS01099E; subsiguientemente, estos productos de amplificación se digirieron con enzimas de restricción. El

modelo obtenido se comparó luego con el modelo que se obtiene cuando dichas digestiones se realizan en productos de amplificación derivados de cepas que contienen genes *cry2A* conocidos. En base al modelo de digestión por restricción, los genes *cry2A* procedentes de las cepas BTS02761A y BTS01099E parecían ser nuevos. Por lo tanto, se secuenció el producto de amplificación. Esto confirmó que los fragmentos amplificados se derivaban de nuevos genes *cry2A*: la cepa BTS02761A contenía un nuevo gen similar a *cry2A*, mientras que la cepa 1099E contenía dos nuevos genes similares a *cry2A*.

ADN total procedente de las cepas BTS02761A y BTS01099E se trató con Sau3A, se fraccionó por tamaños y fragmentos de 7 a 10 kb se ligaron en pUC19I (un derivado de pUC19), se cortaron con BamHI y se trataron TsAP (fosfatasa alcalina térmicamente estable). Esta mezcla de ligamiento se electroporó en XL1 Blue de *E. coli*.

Hibridaciones de las colonias, utilizando los fragmentos de la PCR marcados con DIG como sondas, identificaron clones positivos. Las cepas de *E. coli* recombinantes se depositaron el 6 de octubre de 2000 en la Vakgroep voor Moleculaire Biologie-Plasmidencollectie, Universiteit Gent, K.L. Ledeganckstraat, 35 B-9000 Gent, Bélgica (abreviado aquí en lo que sigue como "BCCM-LMBP") bajo los siguientes números de acceso: BCCM-LMBP 4247 para la cepa XL1Blue:pUC1099E/*cry2clone1*, que codifica una proteína denominada Cry2Af; BCCM-LMBP 4248 para la cepa XL1Blue:pUC1099E/*cry2clone7*, que codifica una proteína denominada Cry2Ae; y BCCM-LMBP 4249 para la cepa XL1Blue:pUC2761A/*cry2clone141*, que codifica una proteína denominada Cry2Ag. Los genes se pueden aislar a partir de estos clones depositados mediante una digestión NotI-Fsel.

El inserto procedente de estos clones se subclonó en el vector de lanzadera pSL40I. El plásmido resultante se transformó primero en GM2163 de *E. coli*. Una preparación de plásmido procedente de esta cepa se electroporó luego en una cepa berliner 1715 de la variedad *B. thuringiensis* cristal-menos.

Un extracto alcalino preparado a partir de una mezcla de esporas/cristales de las cepas de *Bt* recombinantes se utilizó luego en bioensayos para evaluar la toxicidad de las nuevas proteínas Cry2A. Este extracto se testó en el ensayo según se describe antes en el Ejemplo 1. Los resultados se muestran en la Tabla II:

Tabla II:

Toxina	Conc.	Mortalidad (%)					
		Ha	Sf	On	Sn	Hz	Hv
Cry2Ae	1930	83	44	NT	100	100	NT
Cry2Ag	1160	0	0	78	50	29	100
Cry2Aa	470	61	55	50	94	95	100

"Conc.": concentración total de proteínas del extracto de la cepa utilizando el método de Bradford (microg/ml); "Ha": *Heliothis armigera*, las otras abreviaturas son como las utilizadas anteriormente en la Tabla I; los testigos incluidos (dieta normal, adición de PBS-BSA o cepa 1715 de *Bt* cristal-menos no transformada) no proporcionan una mortalidad significativa.

También, el clon recombinante que expresa la proteína Cry2Af muestra una mortalidad significativa cuando se somete a ensayo en insectos lepidópteros seleccionados.

También, se realizó un análisis para determinar los valores LC50 y LC90 para la proteína Cry2Ae producida de forma recombinante, en comparación con las proteínas Cry2Aa y Cry2Ab conocidas.

Para este ensayo, una dieta artificial específica para insectos se dispensó en pocillos de placas Costar de 24 pocillos. 50 microlitros de extracto alcalino (pH 12) de mezclas de esporas-cristales de la cepa *Bt* recombinante que contiene el gen *cry2Ae* que proceden de XL1Blue:pUC1009E/*clone7*, se aplicó sobre la superficie de la dieta y se secó en un flujo de aire laminar. La dieta para *S. frugiperda* y *O. nubilalis* contenía: 1000 ml de agua; agar: 20g; harina de maíz: 112 g; germen de trigo: 28 g; levadura: 30 g; ácido ascórbico: 4,8 g; ácido benzoico: 1,2 g; nipagina: 1 g; aureomicina: 0,06 g; nistatina: 0,03 g. La dieta para *H. virescens* en *H. zea* contenía: 1000 ml de agua; agar: 20 g; harina de soja: 81 g; germen de trigo: 36 g; sacarosa: 14,7 g; aceite de maíz: 5 ml; mezcla de sales de Wesson: 10 g; mezcla de vitaminas Vanderzant: 9,5 g; ácido sórbico: 1,1 g; nipagina: 1 g; aureomicina: 0,34 g; nistatina: 0,06 g. Se sometieron a ensayo diferentes concentraciones de proteínas de modo que se podía determinar un valor LC50. Para los ensayos en *H. zea*, *H. virescens* y *S. frugiperda*, se colocó una larva en cada uno de los pocillos y se utilizaron 20 larvas por cada muestra. Para los ensayos en *O. nubilalis*, se colocaron dos larvas en cada uno de los pocillos y se utilizaron 24 larvas por muestra. Se hizo un recuento el séptimo día de las larvas muertas y vivas (el sexto día para *S. frugiperda* y el quinto día para *O. nubilalis*). Los valores LC50 y LC90 se calcularon con el análisis probit (programa POLO, Software LeOra, 1987, POLO-PC. "A user's guide to probit or logic analysis". Berkeley, California). Los resultados se muestran en la Tabla III que figura a continuación.

Tabla III:

Toxina	Conc.	Valores LC50 (LC90), ambos en ng/cm ²			
		Sf	Hz	Hv	On
Cry2Ae	1160 (*1930)	1154 (3708)	62 (655)	10 (20)	*188 (*1383)
Cry2Aa	2910 (*470)	2906 (10945)	1921 (7740)	35 (138)	*294 (*2854)
Cry2Ab	1290	1498 (8150)	448 (2152)	82 (248)	NT

5 NT: no sometido a ensayo; Conc.: concentración total de proteínas en el extracto alcalino de la cepa de Bt recombinante que produce la proteína relevante en microg/ml; un asterisco indica que el resultado para *O. nubilalis* se obtuvo con una tanda diferente con una concentración diferente de proteínas (indicada entre paréntesis bajo la columna "Conc."); los testigos (dieta normal, PBS-BSA añadido o cepa de Bt testigo cristal-menos) no dieron más de 0-5% de mortalidad.

10 Utilizando la misma estipulación experimental que anteriormente para *Ostrinia nubilalis*, pero utilizando proteína Cry2Ae purificada contra la isoca de las leguminosas, *Anticarsia gemmatalis*, (ensayo de 20 pocillos con 1 larva por concentración), se encontró una elevada actividad de esta proteína contra esta importante plaga de insectos para la soja. Se encontró que el valor LC₅₀ para la proteína Cry2Ae purificada para este insecto era 0,44 ng/cm² (a un nivel de confianza del 95%; este valor LC₅₀ es el valor medio de 2 ensayos de diferentes bio-tandas de proteína purificada), se encontró que el valor LC₉₀ era 7,79 ng/cm² (a un nivel de confianza del 95%; este valor LC₉₀ es el valor medio de 2 bioensayos de diferentes tandas de proteína purificada). Utilizando la misma estipulación experimental que antes para *Ostrinia* con proteína Cry2Ae purificada, se confirmó la importante toxicidad de esta proteína para *Helicoverpa zea* y *Ostrinia nubilalis* (se encontró para estos insectos los valores LC₅₀ de 145,1 y 48,31 ng/cm², respectivamente (a un nivel de confianza del 95%, estos valores LC₅₀ son los valores medios de 2 bioensayos de diferentes tandas de proteína purificada en cada uno de los insectos respectivos)).

25 Estos resultados demuestran que las nuevas proteínas Cry de la invención y, particularmente, la proteína Cry2Ae son proteínas útiles con una elevada actividad contra plagas de insectos lepidópteros relevantes, particularmente contra *Heliothis zea*, *Ostrinia nubilalis*, *Anticarsia gemmatalis* y *Helicoverpa zea*, que son plagas de insectos comercialmente dañinas para plantas tales como soja, algodón y maíz.

30 Las secuencias determinadas para los genes *cry2A* aislados de la invención y las secuencias de aminoácidos determinadas se muestran en el Listado de Secuencias adjunto. Los alineamientos apareados utilizando el programa GAP en el paquete Wisconsin de GCG indicaban los niveles de identidad de la secuencia con otras secuencias de Cry2A (para las secuencias de las proteínas Cry2A conocidas y ADNs, véase Crickmore et al. (1998) y la página web de internet arriba reseñada), tal como se muestra en las Tablas IVA y IVB apareados (valores por defecto de GCG utilizados dentro del programa GAP; para las comparaciones de las secuencias de aminoácidos se utilizó la matriz de registro blosum62, para las comparaciones de secuencias de ADN se utilizó la matriz de registro nwsgapdna).

Tabla IV.A. Porcentaje de identidad de la secuencia al nivel de proteínas:

	Cry2Ae1	Cry2Af1	Cry2Ag1
Cry2Aa1	90,837	88,942	78,905
Cry2Ab1	89,889	94,471	77,331
Cry2Ac1	80,547	80,386	79,869
Cry2Ad1	87,362	91,943	76,849
Cry2Ae1		93,365	79,871
Cry2Af1			79,549

Tabla IV.B. Porcentaje de identidad de la secuencia al nivel de ADN:

	<i>cry2Ae1</i>	<i>cry2Af1</i>	<i>cry2Ag1</i>
<i>cry2Aa1</i>	91,206	89,995	81,994
<i>cry2Ab1</i>	91,890	94,839	81,404
<i>cry2Ac1</i>	84,298	85,209	84,041
<i>cry2Ad1</i>	90,627	93,470	81,136
<i>cry2Ae1</i>		94,576	81,589
<i>cry2Af1</i>			82,233

Ejemplo 4: Producción de las nuevas proteínas Cry en tandas transformadas

5

Genes quiméricos, cada uno de los cuales codifica las proteínas Cry2Ae, Cry2Af y Cry2Ag, se preparan utilizando procesos bien conocidos, utilizando promotores tales como los promotores 35S del CaMV (Hull y Howell, 1987) y de ubiquitina (Christensen et al., 1992). Preferiblemente, la utilización de codones del marco de lectura abierto se adapta al de la planta hospedante con el fin de optimizar la eficacia de la expresión según se describe en la solicitud de patente PCT publicada WO 94/12264. También, en algunos genes quiméricos se incluyen secuencias de ADN que codifican un péptido de tránsito (según se describe en la parte descriptiva) para fijar como objetivo la proteína Cry2A de la invención a los cloroplastos de las plantas.

10

Para la transformación de maíz y algodón con un gen quimérico que codifica la proteína Cry2Ae, varias construcciones de genes quiméricos se insertaron en plásmidos de la cepa de *Agrobacterium*. Estas construcciones incluían: construcciones pACS9 y pACS11, en donde la secuencia codificadora de *cry2Ae* de SEQ ID N° 7 estaba funcionalmente enlazada al promotor 35S2 del virus del mosaico de la coliflor (Odell et al., 1985), una secuencia conductora procedente del gen de proteína de unión a/b de clorofila de *Petunia* (Harpster et al., 1988), y una región de la terminación del transcrito 3' y de poliadenilación del gen 35S procedente del virus del mosaico de la coliflor (Sanfacon et al., 1991), y construcciones pACS12 y pACS13 con las mismas regiones reguladoras y la misma región codificadora de *cry2Ae*, excepto que también una secuencia de ADN que codifica el péptido de tránsito TpssuAt que permite fijar como objetivo los cloroplastos (Krebbes et al., 1988) se insertó en el extremo 5' de la región codificadora de *cry2Ae*, de modo que se produce una proteína de fusión de péptidos de tránsito. Estas construcciones incluían también una secuencia de ADN que codifica una proteína de resistencia a herbicidas glifosato (descrita en la solicitud de patente PCT publicada WO 97/04103, enlazada a un péptido de tránsito optimizado (patente de EE.UU. 5.635.618)) o una secuencia de ADN que codifica una proteína de resistencia al herbicida glufosinato (Thompson et al., 1987) en calidad de marcador seleccionable bajo el control del promotor CsVMV del virus de mosaico de las nervaduras de la mandioca (Verdaguer et al., 1996, 1998) y la región de terminación del transcrito 3' y de poliadenilación del gen nopalina sintasa (Depicker et al., 1982).

15

20

25

Células de maíz se transformaron de manera estable con las construcciones pACS9, pACS11, pACS12 y pACS13 mediante transformación mediada por *Agrobacterium* según se describe en la patente de EE.UU. 6.140.553, incorporada como referencia. Células de plantas de algodón se transformaron de manera estable con las construcciones pACS9, pACS11, pACS12 y pACS13 utilizando el método de transformación descrito en la publicación de patente PCT WO 00/71733, incorporado en esta memoria como referencia. Células de plantas de arroz se transforman de manera estable con el método descrito en la solicitud de patente PCT publicada WO 92/09696. Células de plantas de tabaco se transformaron de manera estable con las construcciones pACS11 y pACS12 utilizando la transformación mediada por *Agrobacterium*, esencialmente según se describe en la patente EP 0 116 718 o Deblaere et al. (1987).

30

35

Las células transformadas y las plántulas regeneradas a partir de las mismas se hacen crecer en medios que contienen los agentes selectivos fosfotricina o glifosato, de modo que la mayoría, si no la totalidad de las plantas regeneradas serán transformadas.

40

Plantas de tabaco, maíz, algodón y arroz transformadas, regeneradas, se seleccionan mediante ELISA de Cry2A, borrones de transferencia Northern y Southern y de acuerdo con la eficacia insecticida y las características agronómicas. Plantas de la progenie con contenido en genes *cry2A* quiméricas muestran una resistencia mejorada a los insectos en comparación con plantas testigos no transformadas con una segregación apropiada de la resistencia a los insectos y el fenotipo transformado. Mediciones de proteínas y ARN muestran que plantas con una resistencia incrementada a los insectos tienen una mayor expresión de la nueva proteína Cry2A en sus células.

45

REFERENCIAS CITADAS

- Adang et al.(1985). *Gene* 36, 289.
 An et al. (1996). *Plant J.* 10, 107.
- 5 Bennetzen Y Hall.(1982).*J. Biol. Chem.* 257, 3026-3031.
 Berhard, K. and Utz, R., "Production of *Bacillus thuringiensis* insecticides for experimental and commercial uses",
 In *Bacillus thuringiensis*, An Environmental Biopesticide: Theory and Practice, págs. 255-267, comps. Entwistle,
 P.F., Cory, J.S., Bailey, M.J. and Higgs, S., John Wiley and Sons, Nueva York (1993).
 Bih et al. (1999), *J. Biol. Chem.* 274, 22884-22894.
- 10 Choi et al., número de acceso a GenBank AF200816 (1999).
 Christensen et al. (1992) *Plant mol. Biol.* 18, 675-689.
 Christou et al. (1990). *Trends Biotechnology* 8, 145.
 Cordera et al. (1994) *The Plant Journal* 6, 141.
 Cornejo et al. (1993) *Plant Mol. Biol.* 23, 567-581.
- 15 Cornelissen et al. (1986) *EMBO J.* 5, 37-40.
 Crickmore et al. (1998) *Microbiol. Mol. Biol Rev.* 62(3), 807-13.
 Datta et al., *Bio/Technology* 8, 736-740 (1990).
 Deblaere et al. (1987) *Methods in Enzymology* 153, 277-292.
 De Pater et al., 1992, *Plant J.* 2, 834-844.
- 20 Depicker et al., 1982, *J. Molec. Appl. Genetics* 1, 561-573.
 Dulmage, H.T., "Production of Bacteria for Biological Control of Insects" in *Biological Control in Crop Production*,
 comp. Paparizas, D.C., Osmun Publishers, Totowa, N.J., USA, págs. 129-141 (1981).
 English et al., *Insect Biochem. Molec. Biol.* 24, 1025-1035 (1994).
 Estruch et al., (1996), *Proc Natl Acad Sci USA* 93, 5389-94.
- 25 Franck et al., *Cell* 21, 285-294 (1980)
 French et al., *Anal.Biochem.* 156, 417-423 (1986).
 French-Constant y Bowen (2000) *Cell Mol Life Sci* 57, 828-33.
 Fromm et al., *Bio/Technology* 8, 833-839 (1990).
 Gardner et al., *Nucleic Acids Research* 9, 2871-2887 (1981)
- 30 Ge et al., *J. Biol. Chem.* 266, 17954-17958 (1991)
 Gielen et al., *EMBO J* 3, 835-845 (1984).
 Gordon-Kamm et al., *The Plant Cell* 2, 603-618 (1990).
 Gould et al., *Plant Physiol.* 95, 426-434 (1991).
 Harpster et al., (1988), *Molecular and General Genetics* 212, 182-190.
- 35 Hesse et al. (1989), *EMBO J.* 8 2453-2461.
 Hinchee et al. (1988) *Bio/Technology* 6, 915.
 Ho et al. (1989). *Gene* 77, 51-59.
 Höfte et al. , *Appl. and Environm. Microbiol.* 54, 2010-2017 (1988)
 Höfte y Whiteley, *Microbiological Review* 53, 242-255 (1989).
- 40 Hull y Howell, *Virology* 86, 482-493 (1987)
 Itakura et al. (1977). *Science* 198, 1056-1063.
 Jansens et al. (1997) *Crop Science* 37, 1616-1624.
 Keil et al. (1986), *Nucl. Acids Res.* 14, 5641-5650.
 Klösgen et al. (1989), *Mol. Gen. Genet.* 217, 155-161.
- 45 Klösgen y Weil (1991), *Mol. Gen. Genet.* 225, 297-304.
 Kota et al. (1999) *Proc. Natl. Acad. Sci. USA* 96, 1840-1845.
 Krebbers et al. (1988) *Plant Molec. Biol.* 11, 745-759.
 Last et al. (1990) *Theor. Appl. Genet.* 81, 581-588.
 Lereclus et al., *Bio/Technology* 10, 418 (1992).
- 50 Mahillon et al, *FEMS Microbiol. Letters* 60, 205-210 (1989).
 Maxam y Gilbert, *Methods in Enzymol.* 65, 499-560 (1980).
 McBride et al., 1995, *Bio/Technology* 13, 362
 Morris et al. (1999), *Biochem. Biophys. Res. Commun.* 255, 328-333.
 Murray et al., M., *Nucleic Acids Research* 17(2), 477-498 (1989).
- 55 Neuhaus y Rogers (1998), *Plant Mol. Biol.* 38, 127-144.
 Odell et al. (1985) *Nature* 313, 810-812.
 Oelmuller et al., *Mol. Gen. Genet.* 237, 261-272 (1993).
 Park et al. (1997), *J. Biol. Chem.* 272, 6876-6881.

- Sanfacon et al. (1991), *Genes and Development* 5, 141-149.
 Sanger et al., *Proc. Natl. Acad. Sci. U S A.* 74(12), 5463-5467 (1977).
 Schnepf et al. (1985). *Journal of Biological Chemistry* 260, 6264.
 Schnepf et al. (1998). *Microbiol. Mol. Biol. Rev.* 62(3), 775-806.
- 5 Shcherban et al. (1995), *Proc. Natl. Acad. Sci USA* 92, 9245-9249.
 Shimamoto et al., *Nature* 338, 274-276 (1989).
 Stanssens et al., *Nucleic Acids Research* 12, 4441-4454 (1989).
 Sutliff et al. (1991) *Plant Molec. Biol.* 16, 579-591.
 Tavladoraki et al. (1998), *FEBS Lett.* 426, 62-66.
- 10 Terashima et al. (1999), *Appl. Microbiol. Biotechnol.* 52, 516-523.
 Thompson et al. (1987), *EMBO J.* 6, 2519-2523.
 Vaeck et al., 1987, *Nature* 328, 33-37.
 Van Den Broeck et al., 1985, *Nature* 313, 358.
 Vanderzant, *J. Econ. Entomol.* 55, p. 140 (1962).
- 15 Van Rie et al., *Science* 247, 72 (1990).
 Velten et al., *J., EMBO J* 3, 2723-2730 (1984).
 Velten y Schell, *Nucleic Acids Research* 13, 6981-6998 (1985)
 Verdaguer et al., *Plant Mol. Biol.* 31, 1129-1139 (1996).
 Verdaguer et al., *Plant Mol. Biol.* 37, 1055-1067 (1998).
- 20 Visser et al., "Domain-Structure Studies of *Bacillus thuringiensis* Crystal Proteins: A Genetic Approach", In *Bacillus thuringiensis, An Environmental Biopesticide: Theory and Practice*, págs. 71-88, comps. Entwistle, P.F., Cory, J.S., Bailey, M.J. and Higgs, S., John Wiley and Sons, Nueva York (1993).
 Wada et al. (1990). *Nucl. Acids Res.* 18, 2367-1411.
 Waterfield et al.(2001) *Trends Microbiol* 9, 185-91.
- 25 White et al.(1989). *Trends in Genet.* 5, 185-189.
 Wong et al.(1992), *Plant Molec. Biol.*20, 81-93.
 Zhang et al. (1991) *The Plant Cell* 3, 1155-1165.

LISTADO DE SECUENCIAS

<110> Aventis CropScience N.V.

5 <120> Nuevas proteínas insecticidas de Bacillus thuringiensis

<130> NEW2AS WO1

<150> US 09/756296

10 <151> 09-01-2001

<160> 9

<170> PatentIn versión 3.0

15 <210> 1

<211> 1899

<212> ADN

<213> Bacillus thuringiensis

20 <220>

<221> CDS

<222> (1) (1896)

<400> 1

atg	aat	aat	gta	tta	aat	aac	gga	aga	act	act	att	tgt	gat	gcg	tat	48
Met	Asn	Asn	Val	Leu	Asn	Asn	Gly	Arg	Thr	Thr	Ile	Cys	Asp	Ala	Tyr	
1				5					10					15		
aat	gta	gtg	gcc	cat	gat	cca	ttt	agt	ttt	gag	cat	aaa	tca	tta	gat	96
Asn	Val	Val	Ala	His	Asp	Pro	Phe	Ser	Phe	Glu	His	Lys	Ser	Leu	Asp	
			20					25					30			
acc	atc	cga	aaa	gaa	tgg	atg	gag	tgg	aaa	aga	aca	gat	cat	agt	tta	144
Thr	Ile	Arg	Lys	Glu	Trp	Met	Glu	Trp	Lys	Arg	Thr	Asp	His	Ser	Leu	
		35					40					45				
tat	gta	gct	cct	ata	gtc	gga	act	gtt	tct	agc	ttt	ctg	cta	aag	aag	192
Tyr	Val	Ala	Pro	Ile	Val	Gly	Thr	Val	Ser	Ser	Phe	Leu	Leu	Lys	Lys	
	50					55					60					
gtg	ggg	agt	ctt	att	gga	aaa	agg	ata	ttg	agt	gaa	tta	tgg	ggg	tta	240
Val	Gly	Ser	Leu	Ile	Gly	Lys	Arg	Ile	Leu	Ser	Glu	Leu	Trp	Gly	Leu	
65					70				75						80	

25

ES 2 397 549 T3

ata	ttt	cct	agt	ggt	agc	aca	aat	cta	atg	caa	gat	att	tta	agg	gag	288
Ile	Phe	Pro	Ser	Gly 85	Ser	Thr	Asn	Leu	Met 90	Gln	Asp	Ile	Leu	Arg 95	Glu	
aca	gaa	caa	ttc	cta	aat	caa	aga	ctt	aat	aca	gac	act	ctt	gcc	cgt	336
Thr	Glu	Gln	Phe 100	Leu	Asn	Gln	Arg	Leu 105	Asn	Thr	Asp	Thr	Leu 110	Ala	Arg	
gta	aat	gcg	gaa	ttg	gaa	ggg	ctg	caa	gcg	aat	ata	agg	gag	ttt	aat	384
Val	Asn	Ala 115	Glu	Leu	Glu	Gly	Leu 120	Gln	Ala	Asn	Ile	Arg 125	Glu	Phe	Asn	
caa	caa	gta	gat	aat	ttt	tta	aat	cct	act	caa	aac	cct	gtt	cct	tta	432
Gln	Gln 130	Val	Asp	Asn	Phe	Leu 135	Asn	Pro	Thr	Gln	Asn 140	Pro	Val	Pro	Leu	
tca	ata	act	tct	tca	gtt	aat	aca	atg	cag	caa	tta	ttt	cta	aat	aga	480
Ser 145	Ile	Thr	Ser	Ser	Val 150	Asn	Thr	Met	Gln	Gln 155	Leu	Phe	Leu	Asn	Arg 160	
tta	ccc	cag	ttc	cgt	gtg	caa	gga	tac	caa	ctg	tta	tta	tta	cct	tta	528
Leu	Pro	Gln	Phe	Arg 165	Val	Gln	Gly	Tyr	Gln 170	Leu	Leu	Leu	Leu	Pro 175	Leu	
ttt	gca	cag	gca	gcc	aat	atg	cat	ctt	tct	ttt	att	aga	gat	gtt	gtt	576
Phe	Ala	Gln	Ala 180	Ala	Asn	Met	His	Leu 185	Ser	Phe	Ile	Arg	Asp 190	Val	Val	
ctc	aat	gca	gat	gaa	tgg	gga	att	tca	gca	gca	aca	tta	cgt	acg	tat	624
Leu	Asn	Ala 195	Asp	Glu	Trp	Gly	Ile 200	Ser	Ala	Ala	Thr	Leu 205	Arg	Thr	Tyr	
caa	aat	tat	ctg	aaa	aat	tat	aca	aca	gag	tac	tct	aat	tat	tgt	ata	672
Gln	Asn 210	Tyr	Leu	Lys	Asn	Tyr 215	Thr	Thr	Glu	Tyr	Ser 220	Asn	Tyr	Cys	Ile	
aat	acg	tat	caa	act	gcg	ttt	aga	ggt	tta	aac	acc	cgt	tta	cac	gat	720
Asn 225	Thr	Tyr	Gln	Thr	Ala 230	Phe	Arg	Gly	Leu	Asn 235	Thr	Arg	Leu	His	Asp 240	
atg	tta	gaa	ttt	aga	aca	tat	atg	ttt	tta	aat	gta	ttt	gaa	tat	gta	768
Met	Leu	Glu	Phe	Arg 245	Thr	Tyr	Met	Phe	Leu 250	Asn	Val	Phe	Glu	Tyr 255	Val	
tct	atc	tgg	tcg	ttg	ttt	aaa	tat	caa	agc	ctt	cta	gta	tct	tct	ggc	816
Ser	Ile	Trp	Ser 260	Leu	Phe	Lys	Tyr	Gln 265	Ser	Leu	Leu	Val	Ser 270	Ser	Gly	
gct	aat	tta	tat	gca	agc	ggt	agt	gga	cca	cag	cag	act	caa	tca	ttt	864
Ala	Asn 275	Leu	Tyr	Ala	Ser	Gly	Ser 280	Gly	Pro	Gln	Gln	Thr 285	Gln	Ser	Phe	
act	tca	caa	gac	tgg	cca	ttt	tta	tat	tct	ctt	ttc	caa	gtt	aat	tca	912
Thr 290	Ser	Gln	Asp	Trp	Pro	Phe 295	Leu	Tyr	Ser	Leu	Phe 300	Gln	Val	Asn	Ser	
aat	tat	gtg	tta	aat	ggc	ttt	agt	ggc	gct	aga	ctt	acg	cag	act	ttc	960
Asn 305	Tyr	Val	Leu	Asn	Gly 310	Phe	Ser	Gly	Ala	Arg 315	Leu	Thr	Gln	Thr	Phe 320	
cct	aat	att	ggt	ggt	tta	cct	ggt	act	act	aca	act	cac	gca	ttg	ctt	1008
Pro	Asn	Ile	Gly	Gly 325	Leu	Pro	Gly	Thr	Thr 330	Thr	Thr	His	Ala	Leu 335	Leu	
gcg	gca	agg	gtc	aat	tac	agt	gga	gga	ggt	tcg	tct	ggt	gat	ata	ggc	1056
Ala	Ala	Arg	Val 340	Asn	Tyr	Ser	Gly	Gly 345	Val	Ser	Ser	Gly	Asp 350	Ile	Gly	
gct	gtg	ttt	aat	caa	aat	ttt	agt	tgt	agc	aca	ttt	ctc	cca	cct	ttg	1104

ES 2 397 549 T3

Ala	Val	Phe	Asn	Gln	Asn	Phe	Ser	Cys	Ser	Thr	Phe	Leu	Pro	Pro	Leu		
		355					360					365					
tta	aca	cca	ttt	gtt	agg	agt	tgg	cta	gat	tca	ggg	tca	gat	cga	ggg	1152	
Leu	Thr	Pro	Phe	Val	Arg	Ser	Trp	Leu	Asp	Ser	Gly	Ser	Asp	Arg	Gly		
		370				375					380						
ggg	gtt	aat	acc	gtt	aca	aat	tgg	caa	aca	gaa	tcg	ttt	gag	tca	act	1200	
Gly	Val	Asn	Thr	Val	Thr	Asn	Trp	Gln	Thr	Glu	Ser	Phe	Glu	Ser	Thr		
					390					395					400		
tta	ggg	tta	agg	tgt	ggg	gct	ttt	aca	gct	cgt	ggg	aat	tca	aac	tat	1248	
Leu	Gly	Leu	Arg	Cys	Gly	Ala	Phe	Thr	Ala	Arg	Gly	Asn	Ser	Asn	Tyr		
				405					410					415			
ttc	cca	gat	tat	ttt	atc	cgt	aat	att	tca	gga	ggt	cct	tta	ggt	ggt	1296	
Phe	Pro	Asp	Tyr	Phe	Ile	Arg	Asn	Ile	Ser	Gly	Val	Pro	Leu	Val	Val		
			420					425					430				
aga	aat	gaa	gat	tta	aga	aga	ccg	tta	cac	tat	aat	gaa	ata	aga	aat	1344	
Arg	Asn	Glu	Asp	Leu	Arg	Arg	Pro	Leu	His	Tyr	Asn	Glu	Ile	Arg	Asn		
		435					440					445					
ata	gaa	agt	cct	tca	gga	aca	cct	ggg	gga	tta	cga	gct	tat	atg	gta	1392	
Ile	Glu	Ser	Pro	Ser	Gly	Thr	Pro	Gly	Gly	Leu	Arg	Ala	Tyr	Met	Val		
	450					455					460						
tct	gtg	cat	aat	aga	aaa	aat	aat	atc	tat	gcc	gtg	cat	gaa	aat	ggg	1440	
Ser	Val	His	Asn	Arg	Lys	Asn	Asn	Ile	Tyr	Ala	Val	His	Glu	Asn	Gly		
					470					475					480		
act	atg	att	cat	tta	gcg	ccg	gaa	gat	tat	aca	gga	ttc	acc	ata	tcg	1488	
Thr	Met	Ile	His	Leu	Ala	Pro	Glu	Asp	Tyr	Thr	Gly	Phe	Thr	Ile	Ser		
				485					490					495			
ccg	ata	cat	gca	act	caa	gtg	aat	aat	caa	acg	cga	aca	ttt	att	tct	1536	
Pro	Ile	His	Ala	Thr	Gln	Val	Asn	Asn	Gln	Thr	Arg	Thr	Phe	Ile	Ser		
			500					505					510				
gaa	aaa	ttt	gga	aat	caa	ggg	gat	tcc	tta	aga	ttt	gaa	caa	agc	aac	1584	
Glu	Lys	Phe	Gly	Asn	Gln	Gly	Asp	Ser	Leu	Arg	Phe	Glu	Gln	Ser	Asn		
		515					520					525					
acg	aca	gca	cgt	tat	aca	ctt	aga	gga	aat	gga	aat	agt	tac	aat	ctt	1632	
Thr	Thr	Ala	Arg	Tyr	Thr	Leu	Arg	Gly	Asn	Gly	Asn	Ser	Tyr	Asn	Leu		
		530				535					540						
tat	tta	aga	gta	tct	tca	cta	gga	aat	tcc	act	att	cga	ggt	act	ata	1680	
Tyr	Leu	Arg	Val	Ser	Ser	Leu	Gly	Asn	Ser	Thr	Ile	Arg	Val	Thr	Ile		
					550				555						560		
aac	ggg	agg	gtt	tat	act	gct	tca	aat	ggt	aat	act	act	aca	aat	aac	1728	
Asn	Gly	Arg	Val	Tyr	Thr	Ala	Ser	Asn	Val	Asn	Thr	Thr	Thr	Asn	Asn		
				565					570					575			
gat	gga	gtt	aat	gat	aat	ggc	gct	cgt	ttt	tta	gat	att	aat	atg	ggg	1776	
Asp	Gly	Val	Asn	Asp	Asn	Gly	Ala	Arg	Phe	Leu	Asp	Ile	Asn	Met	Gly		
			580					585					590				
aat	gta	gta	gca	agt	gat	aat	act	aat	gta	ccg	tta	gat	ata	aat	gtg	1824	
Asn	Val	Val	Ala	Ser	Asp	Asn	Thr	Asn	Val	Pro	Leu	Asp	Ile	Asn	Val		
		595					600					605					
aca	ttt	aac	tcc	ggg	act	caa	ttt	gag	ctt	atg	aat	att	atg	ttt	ggt	1872	
Thr	Phe	Asn	Ser	Gly	Thr	Gln	Phe	Glu	Leu	Met	Asn	Ile	Met	Phe	Val		
	610					615					620						
cca	act	aat	ctt	cca	cca	ata	tat	taa								1899	
Pro	Thr	Asn	Leu	Pro	Pro	Ile	Tyr										

625

630

<210> 2

<211> 632

<212> PRT

<213> *Bacillus thuringiensis*

<400> 2

Met Asn Asn Val Leu Asn Asn Gly Arg Thr Thr Ile Cys Asp Ala Tyr
1 5 10 15

Asn Val Val Ala His Asp Pro Phe Ser Phe Glu His Lys Ser Leu Asp
20 25 30

Thr Ile Arg Lys Glu Trp Met Glu Trp Lys Arg Thr Asp His Ser Leu
35 40 45

Tyr Val Ala Pro Ile Val Gly Thr Val Ser Ser Phe Leu Leu Lys Lys
50 55 60

Val Gly Ser Leu Ile Gly Lys Arg Ile Leu Ser Glu Leu Trp Gly Leu
65 70 75 80

Ile Phe Pro Ser Gly Ser Thr Asn Leu Met Gln Asp Ile Leu Arg Glu
85 90 95

Thr Glu Gln Phe Leu Asn Gln Arg Leu Asn Thr Asp Thr Leu Ala Arg
100 105 110

Val Asn Ala Glu Leu Glu Gly Leu Gln Ala Asn Ile Arg Glu Phe Asn
115 120 125

Gln Gln Val Asp Asn Phe Leu Asn Pro Thr Gln Asn Pro Val Pro Leu
130 135 140

Ser Ile Thr Ser Ser Val Asn Thr Met Gln Gln Leu Phe Leu Asn Arg
145 150 155 160

Leu Pro Gln Phe Arg Val Gln Gly Tyr Gln Leu Leu Leu Leu Pro Leu
165 170 175

Phe Ala Gln Ala Ala Asn Met His Leu Ser Phe Ile Arg Asp Val Val
180 185 190

Leu Asn Ala Asp Glu Trp Gly Ile Ser Ala Ala Thr Leu Arg Thr Tyr
195 200 205

Gln Asn Tyr Leu Lys Asn Tyr Thr Thr Glu Tyr Ser Asn Tyr Cys Ile
210 215 220

ES 2 397 549 T3

Asn Thr Tyr Gln Thr Ala Phe Arg Gly Leu Asn Thr Arg Leu His Asp
 225 230 235 240
 Met Leu Glu Phe Arg Thr Tyr Met Phe Leu Asn Val Phe Glu Tyr Val
 245 250 255
 Ser Ile Trp Ser Leu Phe Lys Tyr Gln Ser Leu Leu Val Ser Ser Gly
 260 265 270
 Ala Asn Leu Tyr Ala Ser Gly Ser Gly Pro Gln Gln Thr Gln Ser Phe
 275 280 285
 Thr Ser Gln Asp Trp Pro Phe Leu Tyr Ser Leu Phe Gln Val Asn Ser
 290 295 300
 Asn Tyr Val Leu Asn Gly Phe Ser Gly Ala Arg Leu Thr Gln Thr Phe
 305 310 315 320
 Pro Asn Ile Gly Gly Leu Pro Gly Thr Thr Thr His Ala Leu Leu
 325 330 335
 Ala Ala Arg Val Asn Tyr Ser Gly Gly Val Ser Ser Gly Asp Ile Gly
 340 345 350
 Ala Val Phe Asn Gln Asn Phe Ser Cys Ser Thr Phe Leu Pro Pro Leu
 355 360 365
 Leu Thr Pro Phe Val Arg Ser Trp Leu Asp Ser Gly Ser Asp Arg Gly
 370 375 380
 Gly Val Asn Thr Val Thr Asn Trp Gln Thr Glu Ser Phe Glu Ser Thr
 385 390 395 400
 Leu Gly Leu Arg Cys Gly Ala Phe Thr Ala Arg Gly Asn Ser Asn Tyr
 405 410 415
 Phe Pro Asp Tyr Phe Ile Arg Asn Ile Ser Gly Val Pro Leu Val Val
 420 425 430
 Arg Asn Glu Asp Leu Arg Arg Pro Leu His Tyr Asn Glu Ile Arg Asn
 435 440 445
 Ile Glu Ser Pro Ser Gly Thr Pro Gly Gly Leu Arg Ala Tyr Met Val
 450 455 460
 Ser Val His Asn Arg Lys Asn Asn Ile Tyr Ala Val His Glu Asn Gly
 465 470 475 480
 Thr Met Ile His Leu Ala Pro Glu Asp Tyr Thr Gly Phe Thr Ile Ser
 485 490 495

ES 2 397 549 T3

Pro Ile His Ala Thr Gln Val Asn Asn Gln Thr Arg Thr Phe Ile Ser
500 505 510

Glu Lys Phe Gly Asn Gln Gly Asp Ser Leu Arg Phe Glu Gln Ser Asn
515 520 525

Thr Thr Ala Arg Tyr Thr Leu Arg Gly Asn Gly Asn Ser Tyr Asn Leu
530 535 540

Tyr Leu Arg Val Ser Ser Leu Gly Asn Ser Thr Ile Arg Val Thr Ile
545 550 555 560

Asn Gly Arg Val Tyr Thr Ala Ser Asn Val Asn Thr Thr Thr Asn Asn
565 570 575

Asp Gly Val Asn Asp Asn Gly Ala Arg Phe Leu Asp Ile Asn Met Gly
580 585 590

Asn Val Val Ala Ser Asp Asn Thr Asn Val Pro Leu Asp Ile Asn Val
595 600 605

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

Pro Thr Asn Leu Pro Pro Ile Tyr
625 630

- <210> 3
- <211> 1899
- <212> ADN
- 5 <213> Bacillus thuringiensis

- <220>
- <221> CDS
- <222> (1) (1896)

10

<400> 3
atg aat agt gta ttg aat agc gga aga act act att tgt gat gcg tat 48
Met Asn Ser Val Leu Asn Ser Gly Arg Thr Thr Ile Cys Asp Ala Tyr
1 5 10 15

aat gta gtg gct cat gat cca ttt agt ttt caa cat aaa tca tta gat 96
Asn Val Val Ala His Asp Pro Phe Ser Phe Gln His Lys Ser Leu Asp
20 25 30

acc ata caa gaa gaa tgg atg gag tgg aaa aaa gat aat cat agt tta 144
Thr Ile Gln Glu Glu Trp Met Glu Trp Lys Lys Asp Asn His Ser Leu
35 40 45

tat gta gat cct att gtt gga act gtg gct agt ttt ctt tta aag aaa 192

ES 2 397 549 T3

Tyr	Val	Asp	Pro	Ile	Val	Gly	Thr	Val	Ala	Ser	Phe	Leu	Leu	Lys	Lys		
	50					55					60						
gtg	ggg	agt	ctt	gtt	gga	aaa	aga	ata	ctg	agt	gag	tta	cgg	aat	tta	240	
Val	Gly	Ser	Leu	Val	Gly	Lys	Arg	Ile	Leu	Ser	Glu	Leu	Arg	Asn	Leu		
65					70					75					80		
ata	ttt	cct	agt	ggc	agt	aca	aat	cta	atg	caa	gat	att	tta	aga	gag	288	
Ile	Phe	Pro	Ser	Gly	Ser	Thr	Asn	Leu	Met	Gln	Asp	Ile	Leu	Arg	Glu		
				85					90					95			
aca	gaa	aaa	ttc	ctg	aat	caa	aga	ctt	aat	aca	gac	act	ctt	gcc	cgt	336	
Thr	Glu	Lys	Phe	Leu	Asn	Gln	Arg	Leu	Asn	Thr	Asp	Thr	Leu	Ala	Arg		
			100					105					110				
gta	aat	gcg	gaa	ttg	aca	ggg	ctg	caa	gca	aat	gta	gaa	gag	ttt	aat	384	
Val	Asn	Ala	Glu	Leu	Thr	Gly	Leu	Gln	Ala	Asn	Val	Glu	Glu	Phe	Asn		
		115					120					125					
cga	caa	gta	gat	aat	ttt	ttg	aac	cct	aac	cga	aat	gct	ggt	cct	tta	432	
Arg	Gln	Val	Asp	Asn	Phe	Leu	Asn	Pro	Asn	Arg	Asn	Ala	Val	Pro	Leu		
	130				135						140						
tca	ata	act	tct	tca	gtt	aat	aca	atg	cag	caa	tta	ttt	cta	aat	aga	480	
Ser	Ile	Thr	Ser	Ser	Val	Asn	Thr	Met	Gln	Gln	Leu	Phe	Leu	Asn	Arg		
					150					155					160		
tta	acc	cag	ttc	cag	atg	caa	gga	tac	caa	ttg	tta	tta	tta	cct	tta	528	
Leu	Thr	Gln	Phe	Gln	Met	Gln	Gly	Tyr	Gln	Leu	Leu	Leu	Leu	Pro	Leu		
				165					170					175			
ttt	gca	cag	gca	gcc	aat	tta	cat	ctt	tct	ttt	att	aga	gat	ggt	att	576	
Phe	Ala	Gln	Ala	Ala	Asn	Leu	His	Leu	Ser	Phe	Ile	Arg	Asp	Val	Ile		
			180					185					190				
ctt	aat	gca	gac	gaa	tgg	gga	att	tca	gca	gca	aca	tta	cgt	acg	tat	624	
Leu	Asn	Ala	Asp	Glu	Trp	Gly	Ile	Ser	Ala	Ala	Thr	Leu	Arg	Thr	Tyr		
		195					200					205					
caa	aat	cac	ctg	aga	aat	tat	aca	aga	gat	tac	tct	aat	tat	tgt	ata	672	
Gln	Asn	His	Leu	Arg	Asn	Tyr	Thr	Arg	Asp	Tyr	Ser	Asn	Tyr	Cys	Ile		
	210					215					220						
aat	acg	tat	caa	act	gcg	ttt	aga	ggt	tta	aac	acc	cgt	tta	cac	gat	720	
Asn	Thr	Tyr	Gln	Thr	Ala	Phe	Arg	Gly	Leu	Asn	Thr	Arg	Leu	His	Asp		
	225				230					235					240		
atg	tta	gaa	ttt	aga	aca	tat	atg	ttt	tta	aat	gta	ttt	gag	tat	gta	768	
Met	Leu	Glu	Phe	Arg	Thr	Tyr	Met	Phe	Leu	Asn	Val	Phe	Glu	Tyr	Val		
				245					250					255			
tct	atc	tgg	tcg	ttg	ttt	aaa	tat	caa	agc	ctt	cta	gtc	tct	tct	ggc	816	
Ser	Ile	Trp	Ser	Leu	Phe	Lys	Tyr	Gln	Ser	Leu	Leu	Val	Ser	Ser	Gly		
			260					265					270				
gct	aat	tta	tat	gca	agt	ggt	agt	gga	cca	cag	cag	acc	caa	tca	ttt	864	
Ala	Asn	Leu	Tyr	Ala	Ser	Gly	Ser	Gly	Pro	Gln	Gln	Thr	Gln	Ser	Phe		
		275				280						285					
act	tca	caa	gac	tgg	cca	ttt	tta	tat	tct	ctt	ttc	caa	ggt	aat	tca	912	
Thr	Ser	Gln	Asp	Trp	Pro	Phe	Leu	Tyr	Ser	Leu	Phe	Gln	Val	Asn	Ser		
	290					295					300						
aat	tat	gtg	tta	aat	ggc	ttt	agt	ggc	gct	aga	ctt	acg	cag	act	ttc	960	
Asn	Tyr	Val	Leu	Asn	Gly	Phe	Ser	Gly	Ala	Arg	Leu	Thr	Gln	Thr	Phe		
	305				310				315						320		
cct	aat	att	gtt	ggt	tta	cct	ggt	act	act	aca	act	cac	gca	ttg	ctt	1008	
Pro	Asn	Ile	Val	Gly	Leu	Pro	Gly	Thr	Thr	Thr	Thr	His	Ala	Leu	Leu		

ES 2 397 549 T3

			325			330			335							
gct	gca	agg	gtc	aat	tac	agt	gga	gga	ggt	tcg	tct	ggt	gat	ata	ggc	1056
Ala	Ala	Arg	Val	Asn	Tyr	Ser	Gly	Gly	Val	Ser	Ser	Gly	Asp	Ile	Gly	
			340					345					350			
gct	gtg	ttt	aat	caa	aat	ttt	agt	tgt	agc	aca	ttt	ctc	cca	cct	ttg	1104
Ala	Val	Phe	Asn	Gln	Asn	Phe	Ser	Cys	Ser	Thr	Phe	Leu	Pro	Pro	Leu	
		355					360					365				
tta	aca	cca	ttt	ggt	agg	agt	tgg	cta	gat	tca	ggt	tca	gat	cgg	ggg	1152
Leu	Thr	Pro	Phe	Val	Arg	Ser	Trp	Leu	Asp	Ser	Gly	Ser	Asp	Arg	Gly	
	370					375					380					
ggg	atc	aat	acc	ggt	acc	aat	tgg	caa	aca	gaa	tcc	ttt	gag	aca	act	1200
Gly	Ile	Asn	Thr	Val	Thr	Asn	Trp	Gln	Thr	Glu	Ser	Phe	Glu	Thr	Thr	
385				390						395					400	
tta	ggt	tta	agg	agt	ggt	gct	ttt	aca	gct	cga	ggt	aat	tca	aac	tat	1248
Leu	Gly	Leu	Arg	Ser	Gly	Ala	Phe	Thr	Ala	Arg	Gly	Asn	Ser	Asn	Tyr	
				405					410					415		
ttc	cca	gat	tat	ttt	atc	cgt	aat	att	tcc	gga	ggt	cct	tta	ggt	ggt	1296
Phe	Pro	Asp	Tyr	Phe	Ile	Arg	Asn	Ile	Ser	Gly	Val	Pro	Leu	Val	Val	
			420					425					430			
aga	aat	gaa	gat	tta	aga	aga	ccg	tta	cac	tat	aat	caa	ata	aga	aat	1344
Arg	Asn	Glu	Asp	Leu	Arg	Arg	Pro	Leu	His	Tyr	Asn	Gln	Ile	Arg	Asn	
		435					440					445				
ata	gaa	agt	cct	tca	gga	aca	cct	ggt	gga	tta	cga	gct	tat	atg	gta	1392
Ile	Glu	Ser	Pro	Ser	Gly	Thr	Pro	Gly	Gly	Leu	Arg	Ala	Tyr	Met	Val	
	450					455					460					
tct	gtg	cat	aac	aga	aaa	aat	aat	atc	tat	gcc	ggt	cat	gaa	aat	ggt	1440
Ser	Val	His	Asn	Arg	Lys	Asn	Asn	Ile	Tyr	Ala	Val	His	Glu	Asn	Gly	
465				470						475					480	
act	atg	att	cat	tta	gcg	ccg	gaa	gat	tat	aca	gga	ttt	act	ata	tcg	1488
Thr	Met	Ile	His	Leu	Ala	Pro	Glu	Asp	Tyr	Thr	Gly	Phe	Thr	Ile	Ser	
				485					490					495		
ccg	ata	cat	gca	act	caa	gtg	aat	aat	caa	acg	cga	aca	ttt	att	tct	1536
Pro	Ile	His	Ala	Thr	Gln	Val	Asn	Asn	Gln	Thr	Arg	Thr	Phe	Ile	Ser	
			500					505					510			
gaa	aaa	ttt	gga	aat	caa	ggt	gat	tcc	tta	aga	ttt	gaa	caa	agc	aac	1584
Glu	Lys	Phe	Gly	Asn	Gln	Gly	Asp	Ser	Leu	Arg	Phe	Glu	Gln	Ser	Asn	
		515					520					525				
acg	aca	gct	cgt	tat	aca	ctt	aga	ggg	aat	gga	aat	agt	tac	aat	ctt	1632
Thr	Thr	Ala	Arg	Tyr	Thr	Leu	Arg	Gly	Asn	Gly	Asn	Ser	Tyr	Asn	Leu	
	530					535					540					
tat	tta	aga	gta	tct	tca	ata	gga	aat	tcc	act	att	cga	ggt	act	ata	1680
Tyr	Leu	Arg	Val	Ser	Ser	Ile	Gly	Asn	Ser	Thr	Ile	Arg	Val	Thr	Ile	
545					550					555					560	
aac	ggt	aga	ggt	tat	act	gct	tca	aat	ggt	aat	act	act	aca	aat	aac	1728
Asn	Gly	Arg	Val	Tyr	Thr	Ala	Ser	Asn	Val	Asn	Thr	Thr	Thr	Asn	Asn	
				565					570					575		
gat	gga	ggt	aat	gat	aat	gga	gct	cgt	ttt	tca	gat	att	aat	att	ggt	1776
Asp	Gly	Val	Asn	Asp	Asn	Gly	Ala	Arg	Phe	Ser	Asp	Ile	Asn	Ile	Gly	
			580					585					590			
aat	gta	gta	gca	agt	gat	aat	act	aat	gta	ccg	tta	gat	ata	aac	gtg	1824
Asn	Val	Val	Ala	Ser	Asp	Asn	Thr	Asn	Val	Pro	Leu	Asp	Ile	Asn	Val	
		595					600					605				

ES 2 397 549 T3

aca tta aat tct ggt act caa ttt gag ctt atg aat att atg ttt gtt 1872
 Thr Leu Asn Ser Gly Thr Gln Phe Glu Leu Met Asn Ile Met Phe Val
 610 615 620

cca act aat atc tca cca ctt tat taa 1899
 Pro Thr Asn Ile Ser Pro Leu Tyr
 625 630

<210> 4

<211> 632

<212> PRT

<213> Bacillus thuringiensis

<400> 4

Met Asn Ser Val Leu Asn Ser Gly Arg Thr Thr Ile Cys Asp Ala Tyr
 1 5 10 15

Asn Val Val Ala His Asp Pro Phe Ser Phe Gln His Lys Ser Leu Asp
 20 25 30

Thr Ile Gln Glu Glu Trp Met Glu Trp Lys Lys Asp Asn His Ser Leu
 35 40 45

Tyr Val Asp Pro Ile Val Gly Thr Val Ala Ser Phe Leu Leu Lys Lys
 50 55 60

Val Gly Ser Leu Val Gly Lys Arg Ile Leu Ser Glu Leu Arg Asn Leu
 65 70 75 80

Ile Phe Pro Ser Gly Ser Thr Asn Leu Met Gln Asp Ile Leu Arg Glu
 85 90 95

Thr Glu Lys Phe Leu Asn Gln Arg Leu Asn Thr Asp Thr Leu Ala Arg
 100 105 110

Val Asn Ala Glu Leu Thr Gly Leu Gln Ala Asn Val Glu Glu Phe Asn
 115 120 125

Arg Gln Val Asp Asn Phe Leu Asn Pro Asn Arg Asn Ala Val Pro Leu
 130 135 140

Ser Ile Thr Ser Ser Val Asn Thr Met Gln Gln Leu Phe Leu Asn Arg
 145 150 155 160

Leu Thr Gln Phe Gln Met Gln Gly Tyr Gln Leu Leu Leu Leu Pro Leu
 165 170 175

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

ES 2 397 549 T3

Leu Asn Ala Asp Glu Trp Gly Ile Ser Ala Ala Thr Leu Arg Thr Tyr
 195 200 205
 Gln Asn His Leu Arg Asn Tyr Thr Arg Asp Tyr Ser Asn Tyr Cys Ile
 210 215 220
 Asn Thr Tyr Gln Thr Ala Phe Arg Gly Leu Asn Thr Arg Leu His Asp
 225 230 235 240
 Met Leu Glu Phe Arg Thr Tyr Met Phe Leu Asn Val Phe Glu Tyr Val
 245 250 255
 Ser Ile Trp Ser Leu Phe Lys Tyr Gln Ser Leu Leu Val Ser Ser Gly
 260 265 270
 Ala Asn Leu Tyr Ala Ser Gly Ser Gly Pro Gln Gln Thr Gln Ser Phe
 275 280 285
 Thr Ser Gln Asp Trp Pro Phe Leu Tyr Ser Leu Phe Gln Val Asn Ser
 290 295 300
 Asn Tyr Val Leu Asn Gly Phe Ser Gly Ala Arg Leu Thr Gln Thr Phe
 305 310 315 320
 Pro Asn Ile Val Gly Leu Pro Gly Thr Thr Thr Thr His Ala Leu Leu
 325 330 335
 Ala Ala Arg Val Asn Tyr Ser Gly Gly Val Ser Ser Gly Asp Ile Gly
 340 345 350
 Ala Val Phe Asn Gln Asn Phe Ser Cys Ser Thr Phe Leu Pro Pro Leu
 355 360 365
 Leu Thr Pro Phe Val Arg Ser Trp Leu Asp Ser Gly Ser Asp Arg Gly
 370 375 380
 Gly Ile Asn Thr Val Thr Asn Trp Gln Thr Glu Ser Phe Glu Thr Thr
 385 390 395 400
 Leu Gly Leu Arg Ser Gly Ala Phe Thr Ala Arg Gly Asn Ser Asn Tyr
 405 410 415
 Phe Pro Asp Tyr Phe Ile Arg Asn Ile Ser Gly Val Pro Leu Val Val
 420 425 430
 Arg Asn Glu Asp Leu Arg Arg Pro Leu His Tyr Asn Gln Ile Arg Asn
 435 440 445
 Ile Glu Ser Pro Ser Gly Thr Pro Gly Gly Leu Arg Ala Tyr Met Val
 450 455 460

ES 2 397 549 T3

Ser Val His Asn Arg Lys Asn Asn Ile Tyr Ala Val His Glu Asn Gly
465 470 475 480

Thr Met Ile His Leu Ala Pro Glu Asp Tyr Thr Gly Phe Thr Ile Ser
485 490 495

Pro Ile His Ala Thr Gln Val Asn Asn Gln Thr Arg Thr Phe Ile Ser
500 505 510

Glu Lys Phe Gly Asn Gln Gly Asp Ser Leu Arg Phe Glu Gln Ser Asn
515 520 525

Thr Thr Ala Arg Tyr Thr Leu Arg Gly Asn Gly Asn Ser Tyr Asn Leu
530 535 540

Tyr Leu Arg Val Ser Ser Ile Gly Asn Ser Thr Ile Arg Val Thr Ile
545 550 555 560

Asn Gly Arg Val Tyr Thr Ala Ser Asn Val Asn Thr Thr Thr Asn Asn
565 570 575

Asp Gly Val Asn Asp Asn Gly Ala Arg Phe Ser Asp Ile Asn Ile Gly
580 585 590

Asn Val Val Ala Ser Asp Asn Thr Asn Val Pro Leu Asp Ile Asn Val
595 600 605

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

Pro Thr Asn Ile Ser Pro Leu Tyr
625 630

- <210> 5
- <211> 1884
- <212> ADN
- 5 <213> Bacillus thuringiensis

- <220>
- <221> CDS
- <222> (1) (1881)

10

<400> 5
atg aat aat gta ttg aat agc gaa aga act act aag tgt ggt gcg tat 48
Met Asn Asn Val Leu Asn Ser Glu Arg Thr Thr Lys Cys Gly Ala Tyr
1 5 10 15

aac gta gtg gct cat gat cca ttc agt ttt gaa cat aaa tca tta gat 96
Asn Val Val Ala His Asp Pro Phe Ser Phe Glu His Lys Ser Leu Asp

ES 2 397 549 T3

	20					25					30					
acc	ata	caa	aaa	gaa	tgg	atg	gag	tgg	aaa	aga	act	gat	cat	agt	tta	144
Thr	Ile	Gln	Lys	Glu	Trp	Met	Glu	Trp	Lys	Arg	Thr	Asp	His	Ser	Leu	
		35					40					45				
tat	gta	tct	cct	att	gta	gga	act	ata	gcc	agt	ttt	ctg	tta	aag	aaa	192
Tyr	Val	Ser	Pro	Ile	Val	Gly	Thr	Ile	Ala	Ser	Phe	Leu	Leu	Lys	Lys	
	50					55					60					
ata	gga	ggg	ctt	ata	gga	aaa	aga	ata	tta	agt	gag	tta	aag	aat	tta	240
Ile	Gly	Gly	Leu	Ile	Gly	Lys	Arg	Ile	Leu	Ser	Glu	Leu	Lys	Asn	Leu	
65					70					75					80	
att	ttt	cct	agt	ggg	agt	ata	gaa	tca	atg	caa	gat	att	tta	aga	ggg	288
Ile	Phe	Pro	Ser	Gly	Ser	Ile	Glu	Ser	Met	Gln	Asp	Ile	Leu	Arg	Gly	
				85					90					95		
gca	gaa	caa	ttt	cta	aat	caa	aga	ctt	gat	gca	gac	acc	ttt	agt	cgt	336
Ala	Glu	Gln	Phe	Leu	Asn	Gln	Arg	Leu	Asp	Ala	Asp	Thr	Phe	Ser	Arg	
			100					105					110			
gta	gaa	gca	gaa	ttg	aga	ggg	ctt	caa	gca	aat	gta	gag	gaa	ttt	aat	384
Val	Glu	Ala	Glu	Leu	Arg	Gly	Leu	Gln	Ala	Asn	Val	Glu	Glu	Phe	Asn	
		115					120					125				
cga	caa	gtg	gac	aat	ttt	tta	aac	cca	aat	caa	aac	cct	gcc	cct	tta	432
Arg	Gln	Val	Asp	Asn	Phe	Leu	Asn	Pro	Asn	Gln	Asn	Pro	Ala	Pro	Leu	
	130					135					140					
gca	ata	att	gat	tcg	gtt	aat	aca	ttg	caa	caa	tta	ttc	cta	agt	aga	480
Ala	Ile	Ile	Asp	Ser	Val	Asn	Thr	Leu	Gln	Gln	Leu	Phe	Leu	Ser	Arg	
145					150					155					160	
tta	ccc	cag	ttc	cag	ata	caa	cgc	tat	cag	cta	tta	tta	tta	cct	tta	528
Leu	Pro	Gln	Phe	Gln	Ile	Gln	Arg	Tyr	Gln	Leu	Leu	Leu	Leu	Pro	Leu	
				165					170					175		
ttt	gca	caa	gca	gcc	aat	tta	cac	ctt	tct	ttt	att	aga	gac	ggt	att	576
Phe	Ala	Gln	Ala	Ala	Asn	Leu	His	Leu	Ser	Phe	Ile	Arg	Asp	Val	Ile	
			180					185					190			
ctt	aat	gca	gat	gaa	tgg	gga	ata	cca	gct	gca	acg	gtg	cgc	aca	tat	624
Leu	Asn	Ala	Asp	Glu	Trp	Gly	Ile	Pro	Ala	Ala	Thr	Val	Arg	Thr	Tyr	
		195				200						205				
aga	gag	cac	cta	caa	aga	tat	aca	cgc	gaa	tac	tcc	aat	tat	tgt	ata	672
Arg	Glu	His	Leu	Gln	Arg	Tyr	Thr	Arg	Glu	Tyr	Ser	Asn	Tyr	Cys	Ile	
	210					215					220					
aat	acg	tat	caa	act	gcg	ttt	aga	ggg	tta	aat	gcc	act	tta	cac	gat	720
Asn	Thr	Tyr	Gln	Thr	Ala	Phe	Arg	Gly	Leu	Asn	Ala	Thr	Leu	His	Asp	
225					230					235					240	
ttt	cta	gaa	ttt	aga	aca	tat	atg	ttt	tta	aat	gta	tta	gac	tat	gta	768
Phe	Leu	Glu	Phe	Arg	Thr	Tyr	Met	Phe	Leu	Asn	Val	Leu	Asp	Tyr	Val	
				245					250					255		
tct	atc	tgg	tcg	ttg	ttt	aaa	tat	cag	agc	ctt	ctg	gta	tcc	tct	ggc	816
Ser	Ile	Trp	Ser	Leu	Phe	Lys	Tyr	Gln	Ser	Leu	Leu	Val	Ser	Ser	Gly	
			260					265					270			
gct	aat	tta	tat	gcg	agt	ggt	agt	gga	gta	aca	aat	aga	caa	tca	ttt	864
Ala	Asn	Leu	Tyr	Ala	Ser	Gly	Ser	Gly	Val	Thr	Asn	Arg	Gln	Ser	Phe	
		275				280						285				
act	gca	caa	gac	tgg	cca	ttt	tta	aat	tct	ctt	ttc	caa	ggt	aat	caa	912
Thr	Ala	Gln	Asp	Trp	Pro	Phe	Leu	Asn	Ser	Leu	Phe	Gln	Val	Asn	Gln	
	290					295					300					

ES 2 397 549 T3

aat Asn 305	tat Tyr	gta Val	tta Leu	aca Thr	ggt Gly 310	atg Met	aat Asn	ggt Gly	tat Tyr	agg Arg 315	tat Tyr	act Thr	tta Leu	agt Ser	tct Ser 320	960
gtt Val	ttt Phe	ggt Gly	aca Thr	aat Asn 325	caa Gln	aca Thr	ata Ile	cat His	tct Ser 330	ggt Val	agg Arg	agt Ser	aat Asn	tat Tyr 335	agg Arg	1008
ggc Gly	ggg Gly	ggt Val	tca Ser 340	tct Ser	ggt Gly	tac Tyr	att Ile	gga Gly 345	ggt Val	aat Asn	ctt Leu	agt Ser	gaa Glu 350	ggt Gly	gac Asp	1056
caa Gln	aat Asn	ttt Phe 355	agt Ser	tgt Cys	agt Ser	aca Thr	ttt Phe 360	ttg Leu	gat Asp	cct Pro	tta Leu	gaa Glu 365	aca Thr	ccg Pro	ttt Phe	1104
att Ile	aga Arg 370	agt Ser	tgg Trp	ctg Leu	gat Asp	tca Ser 375	ggt Gly	agc Ser	gat Asp	gat Asp	ggc Gly 380	ttt Phe	aat Asn	tgg Trp	agt Ser	1152
aca Thr 385	gga Gly	gtc Val	ttt Phe	aca Thr	aca Thr 390	act Thr	att Ile	ggt Gly	tta Leu	cct Pro 395	act Thr	tgt Cys	agc Ser	att Ile	ttt Phe 400	1200
tgg Trp	cct Pro	cgt Arg	ggt Gly	aac Asn 405	tcg Ser	aac Asn	tat Tyr	ttt Phe	cca Pro 410	gat Asp	tat Tyr	ttt Phe	ata Ile	cga Arg 415	aat Asn	1248
att Ile	tct Ser	ggt Gly	gtc Val 420	ggt Val	ggt Gly	cgt Arg	ctt Leu	agg Arg 425	aac Asn	gaa Glu	gat Asp	tta Leu	aga Arg 430	aga Arg	cca Pro	1296
cta Leu	tat Tyr	ttt Phe 435	aat Asn	gag Glu	ata Ile	aga Arg	aat Asn 440	ata Ile	gta Val	gga Gly	aat Asn	aac Asn 445	aat Asn	cca Pro	ccg Pro	1344
gca Ala 450	act Thr	gga Gly	tcg Ser	tta Leu	tca Ser	gtc Val 455	gcc Ala	agc Ser	cta Leu	gtc Val	tct Ser 460	gtg Val	cat His	aac Asn	aga Arg	1392
aaa Lys 465	aat Asn	aat Asn	att Ile	tat Tyr	gct Ala 470	gct Ala	cat His	gaa Glu	aat Asn	ggt Gly 475	act Thr	atg Met	att Ile	cat His	ttg Leu 480	1440
gca Ala	ccg Pro	gaa Glu	gat Asp	tat Tyr 485	aca Thr	ggt Gly	ttc Phe	aca Thr	atg Met 490	tca Ser	cca Pro	ata Ile	cat His	gca Ala 495	act Thr	1488
caa Gln	gta Val	aat Asn	aat Asn 500	caa Gln	aca Thr	cga Arg	aca Thr	ttt Phe 505	att Ile	tcc Ser	gag Glu	aaa Lys	tta Leu 510	gga Gly	aac Asn	1536
caa Gln	ggt Gly	gat Asp 515	tcc Ser	ttg Leu	aga Arg	ttt Phe	gaa Glu 520	caa Gln	aca Thr	aat Asn	aca Thr	acg Thr 525	gct Ala	cga Arg	tac Tyr	1584
aca Thr	ttt Phe 530	aga Arg	ggg Gly	aat Asn	gga Gly	aat Asn 535	agt Ser	tac Tyr	aat Asn	ctt Leu	tat Tyr 540	tta Leu	aga Arg	gta Val	tct Ser	1632
tca Ser 545	cta Leu	gga Gly	aat Asn	tcc Ser	aca Thr 550	att Ile	cga Arg	ggt Val	act Thr	ata Ile 555	aac Asn	ggt Gly	aga Arg	ggt Val	tat Tyr 560	1680
act Thr	ggt Val	tca Ser	aac Asn	gtc Val 565	aat Asn	act Thr	act Thr	aca Thr	aat Asn 570	aac Asn	gat Asp	gga Gly	ggt Val	ggt Val 575	gat Asp	1728

ES 2 397 549 T3

aat ggc gct cgt ttt tca gat att aat ata ggt aat gta gtg gca agt	1776
Asn Gly Ala Arg Phe Ser Asp Ile Asn Ile Gly Asn Val Val Ala Ser	
580 585 590	
gct aat act aat ata cca tta gat ata aat gta aca ttt aac tct ggt	1824
Ala Asn Thr Asn Ile Pro Leu Asp Ile Asn Val Thr Phe Asn Ser Gly	
595 600 605	
acg caa ttt gag ctt atg aat att atg ttt gtt cca act aat att cca	1872
Thr Gln Phe Glu Leu Met Asn Ile Met Phe Val Pro Thr Asn Ile Pro	
610 615 620	
cca att tat taa	1884
Pro Ile Tyr	
625	
<210> 6	
<211> 627	
<212> PRT	
<213> Bacillus thuringiensis	
<400> 6	
Met Asn Asn Val Leu Asn Ser Glu Arg Thr Thr Lys Cys Gly Ala Tyr	
1 5 10	
15	
Asn Val Val Ala His Asp Pro Phe Ser Phe Glu His Lys Ser Leu Asp	
20 25 30	
Thr Ile Gln Lys Glu Trp Met Glu Trp Lys Arg Thr Asp His Ser Leu	
35 40 45	
Tyr Val Ser Pro Ile Val Gly Thr Ile Ala Ser Phe Leu Leu Lys Lys	
50 55 60	
Ile Gly Gly Leu Ile Gly Lys Arg Ile Leu Ser Glu Leu Lys Asn Leu	
65 70 75 80	
Ile Phe Pro Ser Gly Ser Ile Glu Ser Met Gln Asp Ile Leu Arg Gly	
85 90 95	
Ala Glu Gln Phe Leu Asn Gln Arg Leu Asp Ala Asp Thr Phe Ser Arg	
100 105 110	
Val Glu Ala Glu Leu Arg Gly Leu Gln Ala Asn Val Glu Glu Phe Asn	
115 120 125	
Arg Gln Val Asp Asn Phe Leu Asn Pro Asn Gln Asn Pro Ala Pro Leu	
130 135 140	
Ala Ile Ile Asp Ser Val Asn Thr Leu Gln Gln Leu Phe Leu Ser Arg	
145 150 155 160	

ES 2 397 549 T3

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

ES 2 397 549 T3

<222> (3)..(1901)

<400> 7

cc atg gct aac aac gtt ctt aac aac ggt agg act act att tgc gat	47
Met Ala Asn Asn Val Leu Asn Asn Gly Arg Thr Thr Ile Cys Asp	
1 5 10 15	
gca tac aac gtt gtt gct cat gat cct ttc tct ttc gag cat aag tct	95
Ala Tyr Asn Val Val Ala His Asp Pro Phe Ser Phe Glu His Lys Ser	
20 25 30	
ctt gat aca att agg aag gag tgg atg gag tgg aag agg act gat cat	143
Leu Asp Thr Ile Arg Lys Glu Trp Met Glu Trp Lys Arg Thr Asp His	
35 40 45	
tct ctt tac gtt gct cct att gtt ggt act gtt tct tct ttc ctt ctt	191
Ser Leu Tyr Val Ala Pro Ile Val Gly Thr Val Ser Ser Phe Leu Leu	
50 55 60	
aag aag gtt ggt tct ctt atc ggt aag agg atc ctt tct gag ctt tgg	239
Lys Lys Val Gly Ser Leu Ile Gly Lys Arg Ile Leu Ser Glu Leu Trp	
65 70 75	
ggt ctt atc ttc cct tct ggt tct act aac ctt atg caa gat att ctt	287
Gly Leu Ile Phe Pro Ser Gly Ser Thr Asn Leu Met Gln Asp Ile Leu	
80 85 90 95	
agg gag act gaa caa ttc ctt aac cag agg ctt aac act gat act ctt	335
Arg Glu Thr Glu Gln Phe Leu Asn Gln Arg Leu Asn Thr Asp Thr Leu	
100 105 110	
gct agg gtt aac gct gag ctt gag ggt ctt caa gct aac att agg gaa	383
Ala Arg Val Asn Ala Glu Leu Glu Gly Leu Gln Ala Asn Ile Arg Glu	
115 120 125	
ttc aac cag caa gtt gat aac ttc ctt aac cct act caa aac cct gtt	431
Phe Asn Gln Gln Val Asp Asn Phe Leu Asn Pro Thr Gln Asn Pro Val	
130 135 140	
cct ctt tct att act tct tct gtt aac act atg caa caa ctt ttc ctt	479
Pro Leu Ser Ile Thr Ser Ser Val Asn Thr Met Gln Gln Leu Phe Leu	
145 150 155	
aac agg ctt cct caa ttc agg gtt caa ggt tac caa ctt ctt ctt ctt	527
Asn Arg Leu Pro Gln Phe Arg Val Gln Gly Tyr Gln Leu Leu Leu Leu	
160 165 170 175	
cct ctt ttc gct caa gct gct aac atg cac cta agc ttc att agg gat	575
Pro Leu Phe Ala Gln Ala Ala Asn Met His Leu Ser Phe Ile Arg Asp	
180 185 190	
ggt gtt ctt aac gct gat gag tgg ggt att tct gct gct act ctt agg	623
Val Val Leu Asn Ala Asp Glu Trp Gly Ile Ser Ala Ala Thr Leu Arg	
195 200 205	
act tac caa aac tac ctt aag aac tac act act gag tac tct aac tac	671
Thr Tyr Gln Asn Tyr Leu Lys Asn Tyr Thr Thr Glu Tyr Ser Asn Tyr	
210 215 220	
tgc att aac act tac caa act gct ttc agg ggt ctt aac act agg ctt	719
Cys Ile Asn Thr Tyr Gln Thr Ala Phe Arg Gly Leu Asn Thr Arg Leu	
225 230 235	
cat gat atg ctt gag ttc agg act tac atg ttc ctt aac gtt ttc gag	767
His Asp Met Leu Glu Phe Arg Thr Tyr Met Phe Leu Asn Val Phe Glu	
240 245 250 255	

ES 2 397 549 T3

tac	ggt	tct	att	tgg	tct	ctt	ttc	aag	tac	cag	tct	ctt	ctt	ggt	tct	815
Tyr	Val	Ser	Ile	Trp 260	Ser	Leu	Phe	Lys	Tyr 265	Gln	Ser	Leu	Leu	Val	Ser	
tct	ggt	gct	aac	ctt	tac	gct	tct	ggt	tct	ggt	cct	caa	caa	act	caa	863
Ser	Gly	Ala	Asn 275	Leu	Tyr	Ala	Ser	Gly 280	Ser	Gly	Pro	Gln	Gln	Thr	Gln	
tct	ttc	act	tct	caa	gac	tgg	cct	ttc	ctt	tac	tct	ctt	ttc	caa	ggt	911
Ser	Phe	Thr	Ser	Gln	Asp	Trp	Pro 295	Phe	Leu	Tyr	Ser	Leu	Phe	Gln	Val	
aac	tct	aac	tac	ggt	ctt	aac	ggt	ttc	tct	ggt	gct	agg	ctt	act	caa	959
Asn	Ser	Asn	Tyr	Val	Leu	Asn 310	Gly	Phe	Ser	Gly	Ala	Arg	Leu	Thr	Gln	
act	ttc	cct	aac	atc	ggt	ggt	ctt	cct	ggt	act	act	act	act	cat	gct	1007
Thr	Phe	Pro	Asn	Ile	Gly 325	Gly	Leu	Pro	Gly	Thr	Thr	Thr	Thr	His	Ala 335	
ctt	ctt	gct	gct	agg	ggt	aac	tac	tct	ggt	ggt	ggt	tct	tct	ggt	gat	1055
Leu	Leu	Ala	Ala	Arg 340	Val	Asn	Tyr	Ser	Gly 345	Gly	Val	Ser	Ser	Gly	Asp	
atc	ggt	gct	ggt	ttc	aac	cag	aac	ttc	tct	tgc	tct	act	ttc	ctt	cct	1103
Ile	Gly	Ala	Val 355	Phe	Asn	Gln	Asn	Phe 360	Ser	Cys	Ser	Thr	Phe	Leu	Pro	
cct	ctt	ctt	act	cct	ttc	ggt	agg	tct	tgg	ctt	gat	tct	ggt	tct	gat	1151
Pro	Leu	Leu	Thr	Pro	Phe	Val	Arg 375	Ser	Trp	Leu	Asp	Ser	Gly	Ser	Asp	
agg	ggt	ggt	ggt	aac	act	ggt	act	aac	tgg	caa	act	gag	tct	ttc	gag	1199
Arg	Gly	Gly	Val	Asn	Thr	Val 390	Thr	Asn	Trp	Gln	Thr	Glu	Ser	Phe	Glu	
tct	act	ctt	ggt	ctt	agg	tgc	ggt	gct	ttc	act	gct	agg	ggt	aac	tct	1247
Ser	Thr	Leu	Gly	Leu	Arg 405	Cys	Gly	Ala	Phe	Thr 410	Ala	Arg	Gly	Asn	Ser 415	
aac	tac	ttc	cct	gat	tac	ttc	att	agg	aac	att	tct	ggt	ggt	cct	ctt	1295
Asn	Tyr	Phe	Pro	Asp 420	Tyr	Phe	Ile	Arg	Asn 425	Ile	Ser	Gly	Val	Pro	Leu	
ggt	ggt	agg	aac	gag	gat	ctt	agg	agg	cct	ctt	cat	tac	aac	gag	att	1343
Val	Val	Arg	Asn 435	Glu	Asp	Leu	Arg	Arg 440	Pro	Leu	His	Tyr	Asn 445	Glu	Ile	
agg	aac	att	gag	tct	cct	tct	ggt	act	cct	ggt	ggt	ctt	agg	gct	tac	1391
Arg	Asn	Ile 450	Glu	Ser	Pro	Ser	Gly 455	Thr	Pro	Gly	Gly	Leu	Arg	Ala	Tyr	
atg	ggt	tct	ggt	cat	aac	agg	aag	aac	aac	atc	tac	gct	ggt	cat	gag	1439
Met	Val 465	Ser	Val	His	Asn	Arg 470	Lys	Asn	Asn	Ile	Tyr 475	Ala	Val	His	Glu	
aac	ggt	act	atg	att	cat	ctt	gct	cct	gag	gat	tac	acc	ggt	ttc	acc	1487
Asn	Gly	Thr	Met	Ile	His 485	Leu	Ala	Pro	Glu	Asp 490	Tyr	Thr	Gly	Phe	Thr 495	
atc	tcc	ccc	atc	cac	gcc	acc	cag	gtc	aat	aat	cag	acc	agg	acc	ttc	1535
Ile	Ser	Pro	Ile	His 500	Ala	Thr	Gln	Val	Asn 505	Asn	Gln	Thr	Arg	Thr	Phe	
atc	tcc	gag	aag	ttc	ggc	aac	cag	ggc	gac	tcc	ctg	agg	ttc	gag	cag	1583
Ile	Ser	Glu	Lys 515	Phe	Gly	Asn	Gln	Gly 520	Asp	Ser	Leu	Arg	Phe 525	Glu	Gln	
tcc	aac	acc	acc	gcc	agg	tac	acc	ctg	agg	ggc	aac	ggc	aac	tcc	tac	1631

ES 2 397 549 T3

Ser	Asn	Thr	Thr	Ala	Arg	Tyr	Thr	Leu	Arg	Gly	Asn	Gly	Asn	Ser	Tyr	
		530					535					540				
aac	ctg	tac	ctc	agg	gtg	tcc	tcc	ctc	ggc	aac	tcc	acc	atc	agg	gtc	1679
Asn	Leu	Tyr	Leu	Arg	Val	Ser	Ser	Leu	Gly	Asn	Ser	Thr	Ile	Arg	Val	
	545					550					555					
acc	atc	aac	ggc	agg	gtg	tac	acc	gcc	tcc	aac	gtg	aac	acc	acc	acc	1727
Thr	Ile	Asn	Gly	Arg	Val	Tyr	Thr	Ala	Ser	Asn	Val	Asn	Thr	Thr	Thr	
560					565					570					575	
aac	aac	gac	ggc	gtc	aac	gac	aac	ggc	gct	agg	ttc	ctg	gac	atc	aac	1775
Asn	Asn	Asp	Gly	Val	Asn	Asp	Asn	Gly	Ala	Arg	Phe	Leu	Asp	Ile	Asn	
				580					585					590		
atg	ggc	aac	gtc	gtg	gcc	tcc	gac	aac	acc	aac	gtg	ccc	ctg	gac	atc	1823
Met	Gly	Asn	Val	Val	Ala	Ser	Asp	Asn	Thr	Asn	Val	Pro	Leu	Asp	Ile	
			595					600				605				
aac	gtg	aca	ttt	aac	tcc	ggc	acc	cag	ttc	gag	ctg	atg	aac	atc	atg	1871
Asn	Val	Thr	Phe	Asn	Ser	Gly	Thr	Gln	Phe	Glu	Leu	Met	Asn	Ile	Met	
		610				615						620				
ttc	gtg	cca	act	aac	ctc	cca	ccc	atc	tac	tgagctagc						1910
Phe	Val	Pro	Thr	Asn	Leu	Pro	Pro	Ile	Tyr							
	625					630										

<210> 8

<211> 633

<212> PRT

<213> Artificial

<400> 8

Met	Ala	Asn	Asn	Val	Leu	Asn	Asn	Gly	Arg	Thr	Thr	Ile	Cys	Asp	Ala
1				5					10					15	
Tyr	Asn	Val	Val	Ala	His	Asp	Pro	Phe	Ser	Phe	Glu	His	Lys	Ser	Leu
			20					25					30		
Asp	Thr	Ile	Arg	Lys	Glu	Trp	Met	Glu	Trp	Lys	Arg	Thr	Asp	His	Ser
		35					40					45			
Leu	Tyr	Val	Ala	Pro	Ile	Val	Gly	Thr	Val	Ser	Ser	Phe	Leu	Leu	Lys
	50					55					60				
Lys	Val	Gly	Ser	Leu	Ile	Gly	Lys	Arg	Ile	Leu	Ser	Glu	Leu	Trp	Gly
65					70					75					80
Leu	Ile	Phe	Pro	Ser	Gly	Ser	Thr	Asn	Leu	Met	Gln	Asp	Ile	Leu	Arg
				85					90					95	
Glu	Thr	Glu	Gln	Phe	Leu	Asn	Gln	Arg	Leu	Asn	Thr	Asp	Thr	Leu	Ala
			100					105					110		
Arg	Val	Asn	Ala	Glu	Leu	Glu	Gly	Leu	Gln	Ala	Asn	Ile	Arg	Glu	Phe

ES 2 397 549 T3

Thr Leu Gly Leu Arg Cys Gly Ala Phe Thr Ala Arg Gly Asn Ser Asn
 405 410 415

Tyr Phe Pro Asp Tyr Phe Ile Arg Asn Ile Ser Gly Val Pro Leu Val
 420 425 430

Val Arg Asn Glu Asp Leu Arg Arg Pro Leu His Tyr Asn Glu Ile Arg
 435 440 445

Asn Ile Glu Ser Pro Ser Gly Thr Pro Gly Gly Leu Arg Ala Tyr Met
 450 455 460

Val Ser Val His Asn Arg Lys Asn Asn Ile Tyr Ala Val His Glu Asn
 465 470 475 480

Gly Thr Met Ile His Leu Ala Pro Glu Asp Tyr Thr Gly Phe Thr Ile
 485 490 495

Ser Pro Ile His Ala Thr Gln Val Asn Asn Gln Thr Arg Thr Phe Ile
 500 505 510

Ser Glu Lys Phe Gly Asn Gln Gly Asp Ser Leu Arg Phe Glu Gln Ser
 515 520 525

Asn Thr Thr Ala Arg Tyr Thr Leu Arg Gly Asn Gly Asn Ser Tyr Asn
 530 535 540

Leu Tyr Leu Arg Val Ser Ser Leu Gly Asn Ser Thr Ile Arg Val Thr
 545 550 555 560

Ile Asn Gly Arg Val Tyr Thr Ala Ser Asn Val Asn Thr Thr Thr Asn
 565 570 575

Asn Asp Gly Val Asn Asp Asn Gly Ala Arg Phe Leu Asp Ile Asn Met
 580 585 590

Gly Asn Val Val Ala Ser Asp Asn Thr Asn Val Pro Leu Asp Ile Asn
 595 600 605

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

Val Pro Thr Asn Leu Pro Pro Ile Tyr
 625 630

<210> 9
 <211> 1910

ES 2 397 549 T3

<212> ADN
 <213> Artificial

<220>

5 <223> secuencia de ADN de cry2Ae artificial para la expresión en maíz

<400> 9

```

ccatggctaa caacgtgctg aacaacggca ggaccacccat ctgcatgca tacaacgtgg      60
tggcccacga cccattcagc ttcgagcaca agagcctgga caccatccgc aaggagtgga      120
tggagtggaa gcgcaccgac cacagcctgt acgtggcccc tatcgtgggc accgtgagca      180
gcttcctgct gaagaagggt ggagcctga tcggcaagag gatcctgagc gagctgtggg      240
gcctgatctt cccaagcggc agcaccaacc tgatgcagga catcctgagg gagaccgagc      300
agttcctgaa ccagcgcctg aacaccgaca ccctggctcg cgtgaacgcc gagctggagg      360
gcctccaggc caacatcagg gaattcaacc agcaggtgga caacttcctg aaccacaacc      420
agaaccagc gccactgagc atcaccagca gcgtgaacac catgcagcag ctgttcctga      480
accgcctgcc acagttccgc gtgcagggt accagctgct gctgctgcca ctgttcgccc      540
aggctgcca catgcaccta agcttcatcc gcgacgtggt gctgaacgcc gacgagtggg      600
gcatcagcgc tgccaccctg cgcacctacc agaactacct gaagaactac accaccgagt      660
acagcaacta ctgcatcaac acctaccaga ccgccttcag gggcctgaac accaggctgc      720
acgacatgct ggagttccgc acctacatgt tcctgaacgt gttcgagtac gtgagcatct      780
ggagcctggt caagtaccag agcctgctgg tgagcagcgg tgccaacctg tacgccagcg      840
gcagcgggtc acagcagacc cagagcttca ccagccagga ctggcccttc ctgtacagcc      900
tgttccaggt gaacagcaac tacgtgctga acggcttcag cggtgccagg ctgaccagga      960
ccttccaaa catcggaggc ctgccaggca ccaccaccac ccacgccctg ctggctgcca     1020
gggtgaacta cagcgggtggc gtgagcagcg gcgatatcgg cgctgtgttc aaccagaact     1080
tcagctgag caccttcctg ccaccactgc tgacccatt cgtgcgcagc tggctggaca     1140
gcggcagcga caggggtggc gtgaacaccg tgaccaactg gcagaccgag agcttcgaga     1200
gcaccctggg cctgcgctgc ggtgccttca ccgccagggg caacagcaac tacttcccag     1260
actacttcat ccgcaacatc agcggcgtgc cactggtggt gcgcaacgag gacctgcgca     1320
ggccactgca ctacaacgag atccgcaaca tcgagagccc aagcggcacc ccaggaggcc     1380
tgagggccta catggtgagc gtgcacaacc gcaagaacaa catctacgcc gtgcacgaga     1440
acggcaccat gatccacctg gccccagagg actacaccgg tttaccatc tccccatcc     1500
acgccacca ggtcaataat cagaccagga ccttcatctc cgagaagttc ggcaaccagg     1560
gcgactccct gaggttcgag cagtccaaca ccaccgccag gtacaccctg aggggcaacg     1620
gcaactccta caacctgtac ctcaggggtg cctccctcgg caactccacc atcaggggtca     1680
ccatcaacgg caggggtgtac accgcctcca acgtgaacac caccaccaac aacgacggcg     1740
tcaacgacaa cggcgctagg ttctgggaca tcaaatggg caacgctcgtg gcctccgaca     1800
acaccaacgt gccctggac atcaacgtga catttaactc cggcaccag ttcgagctga     1860
    
```


tgaacatcat gttcgtgcca actaacctcc cacccatcta ctgagctagc

1910

REIVINDICACIONES

- 5 1.- Una secuencia de ácidos nucleicos que codifica una proteína que comprende la secuencia de aminoácidos seleccionada del grupo que consiste en: a) la secuencia de aminoácidos de la proteína codificada por el gen *cry2Af* depositada en la BCCM-LMBP bajo del número de acceso LMBP 4247, b) la secuencia de aminoácidos de la proteína de SEQ ID N° 4, c) la secuencia de aminoácidos de SEQ ID N° 4 desde la posición 1 de aminoácido a la posición 625 de aminoácido, y d) la secuencia de aminoácidos de una proteína insecticida codificada por un ADN que se hibrida a la secuencia codificadora de *cry2Af* depositada en la BCCM-LMBP bajo el número de acceso LMBP 4247 bajo condiciones de hibridación rigurosas, teniendo dicho ADN una identidad de la secuencia de al menos el 98% con la secuencia codificadora de SEQ ID N° 3.
- 10 2.- La secuencia de ácidos nucleicos de la reivindicación 1, en donde dicha secuencia de ácidos nucleicos es ADN.
- 15 3.- La secuencia de ácidos nucleicos de la reivindicación 2, que comprende una secuencia de ADN artificial que tiene una utilización de los codones diferente en comparación con la secuencia de ADN que se produce de forma natural, pero que codifica la misma proteína o su fragmento insecticida.
- 20 4.- Una proteína que comprende una secuencia de ácidos nucleicos, seleccionada del grupo que consiste en: a) la secuencia de aminoácidos de la proteína codificada por el gen *cry2Af* depositada en la BCCM-LMBP bajo del número de acceso LMBP 4247, b) la secuencia de aminoácidos de SEQ ID N° 4 desde la posición 1 de aminoácido a la posición 625 de aminoácido, o una proteína de este tipo con 5 a 10 aminoácidos añadidos, reemplazados o suprimidos sin modificar la actividad insecticida de la proteína, o c) la secuencia de aminoácidos de una proteína insecticida codificada por un ADN que se hibrida a la secuencia codificadora de *cry2Af* depositada en la BCCM-LMBP bajo el número de acceso LMBP 4247 bajo condiciones de hibridación rigurosas, teniendo dicha proteína una identidad de la secuencia de al menos el 98% con la proteína de SEQ ID N° 4.
- 25 5.- La proteína de la reivindicación 4, que comprende la secuencia de aminoácidos de SEQ ID N° 4, o una proteína de este tipo con menos de 5 aminoácidos añadidos, reemplazados o suprimidos sin modificar la actividad insecticida de la proteína.
- 30 6.- Un gen quimérico que comprende la secuencia de ácidos nucleicos de la reivindicación 3 o una secuencia de ácidos nucleicos que codifica la proteína de la reivindicación 4 ó 5 bajo el control de un promotor expresable en plantas.
- 35 7.- El agente quimérico de la reivindicación 6 que comprende, además, un ADN que codifica un péptido diana o de tránsito destinado a fijar como objetivo las vacuolas, mitocondrias, cloroplastos, plástidos, o para la secreción.
- 40 8.- Células vegetales, plantas o semillas transformadas de modo que comprendan el gen quimérico de la reivindicación 6.
- 45 9.- Las células vegetales, plantas o semillas de la reivindicación 8, que se seleccionan del grupo que consiste en: maíz, algodón, arroz, tabaco, colza oleaginosa, especies de Brassica, berenjena, soja, patata, girasol, tomate, caña de azúcar, té, habas, tabaco, fresa, clavo, pepino, sandía, pimiento, avena, cebada, trigo, dalia, gladiolo, crisantemo, remolacha azucarera, sorgo, alfalfa y cacahuete.
- 50 10.- Un microorganismo transformado de modo que comprenda la secuencia de ácidos nucleicos de una cualquiera de las reivindicaciones 1 a 3.
- 55 11.- Un procedimiento para hacer a una planta resistente a un insecto, que comprende transformar células vegetales con el gen quimérico de la reivindicación 6, y regenerar plantas transformadas a partir de dichas células que son resistentes a insectos.
- 12.- El procedimiento de la reivindicación 11, en el que dicho insecto se selecciona del grupo que consiste en: *Heliothis virescens*, *Helicoverpa zea*, *Helicoverpa armígera*, *Anticarsia gemmatalis* y *Ostrinia nubilalis*, *Chilo suppressalis*, *Chilo partellus*, *Scirpophaga incertulas*, *Sesamia inferens*, *Cnaphalocrocis medinalis*, *Marasmia patnalis*, *Marasmia exigua*, *Marasmia ruralis* y *Scirpophaga innotata*.
- 13.- Uso de la proteína de la reivindicación 4 ó 5 para reprimir insectos seleccionados del grupo que consiste en:

Heliothis virescens, *Helicoverpa zea*, *Helicoverpa armígera*, *Anticarsia gemmatalis* y *Ostrinia nubilalis*, *Chilo suppressalis*, *Chilo partellus*, *Scirpophaga incertulas*, *Sesamia inferens*, *Cnaphalocrocis medinalis*, *Marasmia patnalis*, *Marasmia exigua*, *Marasmia ruralis* y *Scirpophaga innotata*.