

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 545 683**

21 Número de solicitud: 201430173

51 Int. Cl.:

C07K 14/325 (2006.01)

C12N 15/82 (2006.01)

A01N 63/02 (2006.01)

12

PATENTE DE INVENCION

B1

22 Fecha de presentación:

10.02.2014

43 Fecha de publicación de la solicitud:

14.09.2015

Fecha de la concesión:

30.08.2016

45 Fecha de publicación de la concesión:

06.09.2016

56 Se remite a la solicitud internacional:

PCT/ES2015/070083

73 Titular/es:

**UNIVERSIDAD PÚBLICA DE NAVARRA (55.0%)
Campus de Arrosadia s/n (OTRI) Edificio del
Rectorado
31006 Pamplona (Navarra) ES y
CONSEJO SUPERIOR DE INVESTIGACIONES
CIENTÍFICAS (45.0%)**

72 Inventor/es:

**CABALLERO MURILLO, Primitivo ;
PALMA DOVIS, Leopoldo ;
ESCUDERO FUENTEMILLA, Iñigo Ruíz De ;
MUÑOZ LABIANO, Delia ;
MURILLO MARTINEZ, Jesús y
BERRY, Colin**

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

54 Título: **Nueva proteína Cry de Bacillus thuringiensis con actividad insecticida para combatir un hemíptero.**

57 Resumen:

Nueva proteína Cry de Bacillus thuringiensis con actividad insecticida para combatir un hemíptero.

La invención se refiere a un nuevo gen que codifica una proteína Cry (o delta-endotoxina), a la propia proteína codificada por dicho gen, a los sistemas para la expresión de la misma (vectores y células hospedadoras), y a la utilización de dicho gen y/o dicha proteína para combatir plagas de hemípteros, con preferencia por el pulgón verde, Myzus persicae, tanto mediante el uso de una composición que contenga la nueva proteína y/o un sistema para su expresión, como mediante plantas transgénicas en las que se expresa el gen. Dichas composiciones y plantas transgénicas son de especial interés, pues la nueva proteína Cry presenta una capacidad insecticida con Myzus persicae especialmente buena, en comparación con otras proteínas Cry conocidas, que presentaban una capacidad insecticida reducida contra áfidos y no se conocía ninguna activa con Myzus persicae.

ES 2 545 683 B1

DESCRIPCIÓN

Nueva proteína Cry de *Bacillus thuringiensis* con actividad insecticida para combatir un hemíptero

5

Campo de la invención

La presente invención se refiere al campo de los biopesticidas y de las plantas transgénicas capaces de expresarlos. Más concretamente, la invención se refiere a un nuevo gen que codifica una proteína Cry (o delta-endotoxina), a la propia proteína
10 codificada por dicho gen, a los sistemas para la expresión de la misma (vectores y células hospedadoras), y a la utilización de dicho gen y/o dicha proteína para combatir plagas de *Myzus persicae* (Sulzer), el áfido comúnmente conocido como pulgón verde, tanto mediante el uso de una composición que contenga la proteína como mediante plantas transgénicas en las que se expresa el gen.

15 Antecedentes de la invención

Bacillus thuringiensis (Bt) es una bacteria Gram positiva, ubicua y esporulante que frecuentemente sintetiza diversas proteínas (Cry, Cyt y Vip) con actividad insecticida contra una amplia gama de insectos que causan plagas tales como lepidópteros (Dulmage, 1970; Estruch et al., 1996), coleópteros (Estruch and Yu, 1998; Krieg et al.,
20 1983), dípteros (Federici et al., 2003; Goldberg, 1977) y hemípteros (Porcar et al., 2009; Sattar and Maiti, 2011). Bt produce proteínas insecticidas durante la esporulación que aparecen como cristales paraesporales, cristales que están compuestos predominantemente de una o más proteínas Cry o Cyt. Las proteínas Cry se denominan así porque son las proteínas de las inclusiones paraesporales (cristales, que dan lugar a
25 la abreviatura Cry) de Bt que presentan un efecto tóxico experimentalmente verificable frente a un organismo diana o que tienen una similitud de secuencia significativa con una proteína Cry conocida, mientras que las proteínas Cyt son las proteínas de las inclusiones paraesporales de Bt que exhiben actividad hemolítica (citolítica, que da lugar a la abreviatura Cyt) o que tienen una similitud de secuencia significativa con una
30 proteína Cyt conocida.

Las proteínas Cry pertenecen a una clase de toxinas bacterianas conocidas como toxinas formadoras de poros. Las proteínas Cry de Bt en particular no exhiben actividad insecticida hasta que no han sido ingeridas y solubilizadas en el intestino del insecto, donde la forma ingerida (protoxina) es hidrolizada por proteasas para dar lugar a la

molécula tóxica activa (toxina). Junto con algunas proteínas Cyt, se engloban dentro del grupo de las delta-endotoxinas. Las proteínas Cry generalmente presentan cinco bloques de secuencia conservados, y tres dominios estructurales conservados: dominio I o dominio N, dominio II o dominio M y dominio III o dominio C, que reciben esa denominación por las regiones donde aparecen en la proteína: zona del extremo amino (N), zona media (M) y zona del extremo carboxilo (C). Esos tres dominios estructurales son responsables de su actividad insecticida contra distintas especies de insectos, así como también contra células cancerígenas de origen humano (parasporinas) (Ohba et al., 2009). De estos últimos, se considera que los dominios M y C en particular están implicados en el reconocimiento y unión al receptor, y se consideran por ello determinantes de su especificidad como toxinas. El dominio III (dominio C) presenta a menudo una homología estructural con proteínas que se unen a hidratos de carbono como el dominio de unión a celulosa de la 1,4- β -glucanasa C, galactosa oxidasa, sialidasa, β -glucoronidasa, el dominio de unión a hidratos de carbono de la xilanasasa U y β -galactosidasa. La forma de protoxina de muchas proteínas Cry presenta una extensión en el extremo carboxilo que hace que su longitud sea mayor que la de otras protoxinas de la familia, extensión que se considera dispensable para la toxicidad pero que se cree que juega un papel en la formación de los cuerpos de inclusión cristalinos dentro de la bacteria.

Los insecticidas basados en delta-endotoxinas han tenido un uso intensivo, que ha dado lugar incluso a la aparición de resistencias en algunas especies.

Dentro del orden de los hemípteros (*Hemiptera*), la familia de los afidos o áfidos (*Aphididae*), más conocidos como pulgones, contiene gran número de parásitos de plantas angiospermas, que originan algunas de las principales plagas que afectan a diferentes cultivos agrícolas. El pulgón verde, *Myzus persicae* (Sulzer, 1776), es una de las principales plagas polífagas en muchas regiones del mundo (Harrington, 2007) incluida España (Díaz et al., 2006). Afecta también a un gran número de cultivos agrícolas de gran importancia económica (Harrington, 2007), entre los que se incluyen distintos frutales (como los de algunos cítricos, el melocotonero y otros árboles del género *Prunus*) y cultivos herbáceos como solanáceas, gramíneas y crucíferas. En dichos cultivos no sólo produce daños directos por alimentación, sino también por ser un eficiente vector transmisor de varios virus fitopatógenos causantes de enfermedades vegetales muy perjudiciales (Nebreda et al., 2004). Es más, recientemente se ha descrito que esta especie se ha hecho resistente a algunos de los insecticidas químicos de síntesis más comúnmente utilizados para su control en Europa (Slater et al., 2012).

Actualmente, se conocen varias proteínas Cry de Bt con actividad insecticida contra áfidos. Sin embargo, no han resultado especialmente tóxicas contra estos insectos, ya que son necesarias elevadas dosis para matar entre el 50% y el 100% de los insectos tratados (Chougule and Bonning, 2012; Porcar et al., 2009). La Tabla 1 que se muestra a continuación resume las toxicidades conocidas de proteínas Cry contra especies de áfidos que causan plagas, de acuerdo con los datos de Chougule and Bonning (Chougule and Bonning, 2012).

Tabla 1

Toxicidades conocidas de proteínas Cry contra áfidos causantes de plagas

Toxina	Toxicidad	Especificidad
Cry2, Cry3, Cry4	Reducida	Pulgón de la patata (<i>Macrosiphum euphorbiae</i>)
Cry4Aa	LC50: 70 – 100 µg/ml	Pulgón verde de la alfalfa (<i>Acyrtosiphon pisum</i>)
Cry11Aa	100% mortalidad a 500 µg/ml	
Cry3A	60% mortalidad a 500 µg/ml	

10

La solicitud de patente internacional WO2010099365 (pertenece a la misma familia que la patente de EE.UU. US8318900B2), se refiere a secuencias nucleotídicas que codifican proteínas que presentan homología con las delta-endotoxinas, a composiciones que comprenden dichas proteínas y a métodos para conferir resistencia contra plagas a bacterias, plantas, células vegetales, tejidos y semillas mediante la transformación de dichos organismos con dichas secuencias nucleotídicas. La solicitud WO2010099365 divulga tanto secuencias naturales de genes identificados en distintas cepas de *Bacillus thuringiensis* (SEQ ID NOs:1 a 47 de dicha solicitud) como secuencias sintéticas preparadas a partir de las anteriores (SEQ ID Nos:97 a 196 de dicha solicitud WO2010099365), que dan lugar a proteínas que carecen del "dominio cristalino" C-terminal presente en la mayor parte de las delta-endotoxinas. En la solicitud WO2010099365 se cita *Myzus persicae* como una de las posibles plagas a combatir. Sin embargo, la solicitud no menciona ningún ensayo en el que se pruebe la eficiencia para combatir dicho áfido de las composiciones y métodos divulgados en el documento WO2010099365, aunque sí se definen unas pautas generales para posibles ensayos de áfidos en el Ejemplo 20, sin mostrar preferencia por el uso de ninguna de las delta-endotoxinas divulgadas en el documento. El Ejemplo 19 de dicha solicitud describe un

25

ensayo en el que se prueba la eficacia contra otro áfido, *Aphis glycines*, de proteínas sintéticas derivadas de la proteína Axmi171 (un homólogo lejano de la proteína Cry12Aa2, codificado por la secuencia insertada en la construcción descrita en la secuencia número 204 de la solicitud WO2010099365 y de la patente US8318900B2, 5 secuencia que codifica la secuencia de aminoácidos identificada con el número 205); el máximo valor de mortalidad obtenido, 2 (es decir, 25-50% de los recipientes mostraban 100% de mortalidad) sólo se consigue con un extracto concentrado de *E. coli* que expresa la forma de la proteína que carece del dominio N-terminal o con una forma de la proteína obtenida tras escindir con el factor Xa una proteína de fusión que incluía la 10 proteína de unión a la maltosa y resuspender en agua el precipitado formado, mientras que el resto de las variantes dieron lugar a 0-25% de los recipientes con 100% de mortalidad. Sin embargo, en el ejemplo anterior, el Ejemplo 18, se describen ensayos realizados también con derivados de dicha proteína Axmi171, concretamente con distintas concentraciones de la forma de la proteína obtenida tras escindir con el factor Xa 15 una proteína de fusión que incluía la proteína de unión a la maltosa y resuspender en agua el precipitado formado, pero donde se determina la actividad contra un hemíptero que no pertenece a la familia de los áfidos sino a otra familia, la familia *Miridae*, concretamente *Lygus hesperus*, encontrando que una concentración de 50 ppm (50 µg/ml) es suficiente para dar lugar a un 80% de mortalidad en dicho hemíptero.

20 La solicitud de patente internacional WO2009158470, perteneciente a la familia de la patente de EE.UU. US8461421B2 y del mismo solicitante que la anterior, describe también nuevos genes que codifican proteínas pesticidas de tipo delta-endotoxinas. Las proteínas y sus respectivos genes se utilizan para preparar formulaciones plaguicidas y para la producción de plantas transgénicas resistentes a las plagas. *Myzus persicae* es 25 mencionado como una de las posibles plagas a combatir mediante la invención descrita en dicha solicitud de patente internacional, sin precisar la posible delta-endotoxina a utilizar o las condiciones concretas de aplicación. En dicha solicitud, no se menciona tampoco ningún ejemplo en el que se demuestre la utilidad de dichas formulaciones o de dichas plantas transgénicas contra ningún áfido, sino únicamente contra algunos 30 lepidópteros, aunque sí se definen unas pautas generales para posibles ensayos de áfidos en el Ejemplo 8, de nuevo sin mostrar preferencia por el uso de ninguna de las delta-endotoxinas divulgadas en el documento para efectivamente llevarlos a cabo.

Las solicitudes de patente de EE.UU. publicadas como US2012047607A1 y US201204760A1 (ambas con el mismo solicitante, Pioneer Hi-Bred International Inc.) se refieren también a nuevos genes de *Bacillus thuringiensis* que codifican proteínas plaguicidas de la familia Cry, a composiciones derivadas de los mismos y a métodos e control de plagas que las emplean, incluida la generación de plantas transgénicas que los expresan. Aunque se citan varios áfidos como posibles plagas de interés, entre ellos *Myzus persicae* en particular, las solicitudes no incluyen ensayos en los que se demuestre la efectividad de las proteínas de dichas invenciones contra ningún áfido. En la solicitud US2012047605A1, por ejemplo, sólo se describen ensayos realizados con varios lepidópteros, alimentados con una dieta artificial que incluye la proteína Cry divulgada en dicha solicitud, obtenida por mutagénesis. En la solicitud US2012047607A1, por su parte, se describe un ensayo similar con un coleóptero, *Diabrotica virgifera*, (Ejemplo 1), que incluye también el cálculo de la LC₅₀ y de la EC₅₀ (concentración efectiva media) (Ejemplo 2).

La solicitud de patente internacional WO9516778 sí se refiere específicamente a un método para combatir los daños que los áfidos causan en plantas mediante una proteína tóxica de *Bacillus thuringiensis*. Dicha solicitud de patente internacional menciona en los antecedentes de la invención la falta de actividad contra los áfidos de los insecticidas basados en *Bacillus thuringiensis* cuando se utilizan de forma independiente, no dando lugar tampoco a un aumento de la efectividad cuando se combinan con piretrinas. La misma falta de actividad se había encontrado con diversas plantas transgénicas capaces de expresar proteínas insecticidas de *Bacillus thuringiensis*. La solicitud WO9516778 propone alimentar áfidos con una dieta artificial que contiene proteínas Cry preparadas en una formulación adecuada (solubilizada en un medio acuoso o en forma de suspensión de sólidos) por medio de un dispositivo alimentador, preparado a escala de laboratorio, donde los áfidos están separados de la dieta por una fina membrana de plástico. Como alternativa se propone la obtención de plantas transgénicas en las que el promotor al que está unido el gen de la proteína Cry se elige para que dicha proteína Cry se exprese en el tejido vascular de la planta, preferiblemente en el floema, pues eso asegura que los áfidos ingieran la proteína al alimentarse de la planta. Dicha solicitud no muestra ningún ejemplo en el que se demuestre la efectividad de la estrategia de expresión de proteínas Cry en el floema de plantas transgénicas.

En cuanto a la alimentación artificial, en la solicitud WO9516778 sólo se describen ensayos con el pulgón de la patata, *Macrosiphum euphorbiae* (Thomas), que demuestran que son necesarias concentraciones elevadas de la proteína CryIIIa para observar un

efecto (Ejemplo 1). En el Ejemplo 2, por su parte, intenta comparar la efectividad de las proteínas CryIA(c), CryIIA, CryIIIA, CryIIIB2, CryIIIB3, CryIVD y de una mezcla de proteínas Cry de tipo I (CryIA(a), CryIA(b), CryIC y CryIF), todas ellas en suspensión acuosa; aunque la diferencia en las concentraciones añadidas en cada caso dificulta las comparaciones, el resultado demuestra que sólo la proteína CryIA(c) no mostró actividad insecticida contra *Macrosiphum euphorbiae*, mientras que la proteína CryIIA (a 200 ng/μl) dio lugar a una rápida aparición de la mortalidad, que era cercana al 100% tras dos días de alimentación, siendo menos eficaces las otras proteínas. En la sección dedicada a la descripción detallada de las realizaciones preferidas de la invención de la solicitud WO9516778, se considera que las proteínas más eficaces para combatir áfidos, junto con la anterior, son las proteínas CryIIIA, CryIIIB y CryIVD. Esta valoración, sin embargo, no ha sido realizada en ensayos con *Myzus persicae*. Las restantes proteínas ensayadas mostraron porcentajes variables de actividad insecticida, necesitándose de 4 a 5 días para observar una mortalidad cercana al 100% en el caso de la mezcla de proteínas Cry (donde la concentración total era próxima a 400 μg/μl, equivalente a 400 mg/ml), en el caso de la proteína CryIIIA (a 400 ng/μl, equivalente a 400 μg/ml) y en el de la proteína CryIVD (a 350 ng/μl, equivalente a 350 μg/ml) y hasta 7 días o más en el caso de las proteínas CryIIIB2 (a 150 ng/μl, equivalente a 150 μg/ml) y CryIIIB3 (132 μg/μl, equivalente a 132 mg/ml).

De acuerdo con lo anterior, puede decirse que no se ha descrito hasta ahora ninguna proteína Cry que posea actividad insecticida contra *Myzus persicae*.

Dada la extensión mundial de *Myzus persicae* y el gran número de cultivos agrícolas a los que afecta, sería interesante encontrar una proteína con actividad insecticida, por ejemplo del tipo de las endotoxinas, que mostrara una actividad insecticida significativa contra áfidos y, en particular, como *Myzus persicae*. La presente invención proporciona una solución a dicho problema.

Breve descripción de la invención

La presente invención se refiere, en un primer aspecto, a una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica una delta-endotoxina, caracterizada por que la secuencia de nucleótidos que codifica una delta-endotoxina está representada por SEQ ID NO:1 o presenta al menos un 90% de identidad con la misma. A la secuencia de nucleótidos que codifica una delta-endotoxina y que está representada por SEQ ID NO:1 se aludirá también en lo sucesivo como "el gen de la presente invención", "el gen *cry* de la presente invención" o "el nuevo gen *cry*". A la molécula de

ácido nucleico definida en este primer aspecto de la invención se aludirá como la molécula de ácido nucleico de la presente invención.

En aspectos adicionales de la invención, la misma se refiere a un sistema de expresión de la molécula de ácido nucleico anterior y/o de la secuencia codificante comprendida en la misma, y a los vectores de expresión que pueden formar parte de dicho sistema de expresión. Por tanto, la presente invención se refiere también a un vector en el que está inserta la molécula de ácido nucleico de la presente invención, vector que se considerará en adelante un vector de la presente invención. Complementariamente, la invención además se refiere a una célula hospedadora que comprende un vector de la invención, es decir, una célula hospedadora de la presente invención.

En un aspecto más, la invención se refiere a una planta transgénica que comprende una célula hospedadora de la presente invención que es una célula de planta, así como a una semilla transgénica que comprende una molécula de ácido nucleico de la presente invención.

Otro aspecto de la invención es un polipéptido de la presente invención, es decir, un polipéptido que comprende una secuencia de aminoácidos caracterizada por que dicha secuencia de aminoácidos:

- a) es la representada por SEQ ID NO:2;
- b) es la secuencia de una delta-endotoxina que está codificada por la secuencia representada por SEQ ID NO:1 o por una secuencia que presenta al menos un 90% de identidad con la misma, o
- c) es la secuencia de una delta-endotoxina que presenta al menos un 80% de identidad con la secuencia representada por SEQ ID NO:2.

La secuencia de aminoácidos representada por SEQ ID NO:2, así como la secuencia que presenta al menos un 80% de identidad con la misma o la secuencia de una delta-endotoxina que está codificada por la secuencia representada por SEQ ID NO:1 o por una secuencia que presenta al menos un 90% de identidad con la misma, se considerarán en adelante la delta-endotoxina de la presente invención.

Es también un aspecto de la invención una composición que comprende un polipéptido de la presente invención, un vector de la presente invención y/o una célula hospedadora de la presente invención. Dichas composiciones se considerarán las composiciones de la presente invención.

También es un aspecto de la invención el uso de una composición de la presente invención para combatir plagas de al menos un hemíptero, con preferencia por que el hemíptero sea un áfido, y especial preferencia por el áfido *Myzus persicae*.

Un aspecto más de la invención es un método para la producción de un polipéptido de la presente invención, que comprende:

- a) obtener un organismo recombinante que exprese el polipéptido de la presente invención;
- b) cultivar dicho organismo recombinante;
- c) inducir, de ser necesario, la expresión del polipéptido en dicho organismo recombinante;
- d) purificar el polipéptido a partir del organismo recombinante.

Finalmente, también es un aspecto de la invención un anticuerpo que se une específicamente a un polipéptido de la presente invención, especialmente un anticuerpo que se une específicamente al fragmento de dicho polipéptido que constituye la delta-endotoxina de la presente invención.

Breve Descripción de las Figuras

Fig. 1: Alineamiento múltiple de secuencias de proteínas Cry activas contra diferentes especies de áfidos. Se incluye la proteína Cry de la presente invención (CryBM1) y la proteína representada por la secuencia con número de identificación 205 en la patente US8318900B2 (Seq205-US8318900). Bajo cada secuencia aparecen rectángulos sombreados que indican, de haberlos, la situación de los dominios N, M y C, característicos de las delta-endotoxinas insecticidas, así como la de otras regiones significativas (Gal: dominio de unión a galactosa; CW: unión a la pared celular; TMR: región transmembrana; SP: péptido señal) Nótese la presencia de un rectángulo con la etiqueta "Ricina B" (que indica un dominio tipo ricina B) bajo el extremo derecho de la representación correspondiente a la proteína CryBM1.

Fig. 2: Dendrograma representativo de la relación filogenética entre proteínas Cry activas contra áfidos. Se incluye la proteína Cry de la presente invención (CryBM1) y la proteína representada por la secuencia con número de identificación 205 en la patente US8318900B2 (Seq205-US8318900). La longitud de la barra horizontal situada bajo la línea de la proteína inferior, Cry1Aa, corresponde a una distancia taxonómica de 0,2,

calculada en base al porcentaje de identidad de los alineamientos de las secuencias de aminoácidos, según se indica en la propia figura.

Fig. 3: Esquema del vector recombinante de expresión pET-28b(+) utilizado para insertar el gen *cryBM1* y utilizado entonces para transformar células de *Escherichia coli* BL21(DE3) y expresar la proteína insecticida de la presente invención bajo control del promotor del bacteriofago T7 (Promotor T7) inducible por IPTG. Se indica la posición de inserción del gen *cryBM1* y el tamaño del plásmido antes de la inserción de dicho gen (5369 pares de bases, abreviado como bp).

Fig. 4: Fotografía del gel de SDS-PAGE obtenido al someter a electroforesis una muestra de la proteína CryBM1 expresada en *Escherichia coli* y purificada en columna de níquel. Se indica con una flecha y el tamaño estimado (89 kDa) la banda correspondiente a la muestra (carril derecho, etiquetado como "Cry"); en el carril izquierdo, correspondiente al marcador, se indica el tamaño de las dos bandas entre las que se localiza la banda de la muestra de CryBM1.

Fig. 5: Fotografía de los elementos utilizados en el bioensayo de actividad insecticida. De izquierda a derecha: vista posterior de un fragmento de las tiras de película plástica para laboratorio Parafilm® "M"; vista lateral de una caja de ensayo; vista superior de una caja de ensayo, con 15 ninfas neonatas de *Myzus persicae*; vista superior de una caja cubierta con una capa de Parafilm® "M" con una gota de dieta + proteína CryBM1 ("toxina" en la leyenda de la fotografía) depositada sobre la misma; vista superior de una caja con los mismos elementos que la anterior, en la que la gota de dieta + proteína CryBM1 se ha cubierto con una segunda capa de Parafilm® "M". Junto a las cajas (en un nivel superior de la superficie de la fotografía) se muestra un extremo de la pipeta automática utilizada para depositar las gotas de dieta + proteína CryBM1, en la que se ha encajado la punta utilizada para tomar y depositar las muestras.

Descripción detallada de la invención

La presente invención se basa en la identificación y aislamiento de un gen de *Bacillus thuringiensis* (Bt), así como en la identificación, aislamiento, producción y evaluación insecticida de su producto, una nueva proteína Cry de Bt que es tóxica contra el áfido *Myzus persicae* o pulgón verde. Este nuevo gen y su producto, pueden ser utilizados para el desarrollo de nuevos biopesticidas y la construcción de plantas transgénicas resistentes al pulgón verde.

En la presente solicitud se detallan la secuencia de nucleótidos del nuevo gen aislado (representada por SEQ ID NO:1) y la secuencia de aminoácidos de la proteína codificada por dicho gen (representada por SEQ ID NO:2). También se aportan pruebas de la capacidad insecticida de dicha proteína, como el cálculo del valor de la concentración letal media (CL₅₀), (que es la concentración de toxina que produce la muerte en el 50% de los insectos que han sido tratados), que ha resultado ser una concentración más baja que las descritas para otras proteínas Cry aplicadas a otros hemípteros plaga, particularmente más baja que las descritas hasta ahora para otros áfidos.

La presente solicitud proporciona una descripción detallada de la nueva proteína Cry y del gen que la codifica, que son significativamente diferentes, en secuencia y en algunas características estructurales, de las secuencias correspondientes a otras proteínas Cry y, en particular, bastante diferentes de las proteínas Cry con actividad insecticida comprobada descritas hasta ahora.

El análisis estructural esquematizado en la Fig. 1 muestra que la proteína cuyo aislamiento y producción se describe más adelante en la presente solicitud (indicada en dicha Fig. 1 como CryBM1) es una proteína de 89 kDa que presenta la estructura característica de las delta-endotoxinas, incluidos los dominios N, M y C que generalmente están presentes en las proteínas Cry. En el dominio C en particular se observa una zona similar a un dominio de unión a hidratos de carbono, concretamente un dominio de unión a galactosa. Por ello, a la proteína cuyo aislamiento y producción se divulga en la presente solicitud, junto con la identificación y aislamiento de su secuencia codificante, se le denomina en la presente solicitud la proteína Cry de la presente invención, la delta-endotoxina de la presente invención o, simplemente, la proteína de la presente invención.

El dendrograma representativo de la relación filogenética entre proteínas Cry activas contra áfidos de la Fig. 2 muestra que la proteína Cry de la presente invención no está estrechamente relacionada con ninguna de dichas proteínas, indicando que se corresponde con una proteína no descrita hasta la fecha. A dicha proteína se le ha asignado provisionalmente la denominación abreviada CryBM1 en relación al grupo de investigación (Bioinsecticidas Microbianos).

Los análisis comparativos realizados, basados en las búsquedas en bases de datos de otras secuencias de nucleótidos conocidas y su comparación con respecto a SEQ ID NO:1, por una parte, y de secuencias de proteínas conocidas y su comparación con respecto a SEQ ID NO:2, por otra parte, han permitido corroborar no sólo que la nueva proteína Cry (o delta-endotoxina) de la presente invención es nueva, como también lo es

su secuencia codificante, por no coincidir sus secuencias con ninguna otra secuencia de aminoácidos o nucleótidos, respectivamente, previamente conocidas, sino también detectar diferencias con otras secuencias relacionadas.

Así, por ejemplo, la comparación de la secuencia de nucleótidos del gen de la presente
5 invención (SEQ ID NO:1) con otras secuencias de nucleótidos ha permitido constatar que
no se encuentra ninguna secuencia de nucleótidos que presente un porcentaje de
identidad que alcance el valor del 90%, o superior, con respecto a SEQ ID NO:1. El
máximo porcentaje de identidad, 87%, corresponde a la secuencia de nucleótidos con el
10 número 38 en la familia de la solicitud de patente internacional WO 2010099365 y la
patente de EE.UU. US8318900, a la que también pertenece la solicitud japonesa JP
2012519000-A, secuencia que codifica la proteína que responde a la secuencia 87 de
dicha familia. También aparecen entre la lista de secuencias con mayor porcentaje de
homología respecto a SEQ ID NO:1 (véase la Tabla 4 que se presenta más adelante)
15 otras secuencias de nucleótidos divulgadas en la misma familia de documentos de
patente, como la secuencia 169 de WO2010099365 (secuencia 122 de JP 2012519000-
A), con la que el gen de la presente invención presenta una identidad del 70%. Aparece
también en esa misma lista una secuencia divulgada en la familia de documentos de
patente de la solicitud internacional WO 2009158470 y la patente de EE.UU. US8461421
(que incluye también a la solicitud de patente de Japón JP2011526150-A), la secuencia
20 con el número 24 de esta segunda familia de patentes, con la que SEQ ID NO:1 de la
presente invención presenta un 71% de identidad (considerando los nucleótidos 1 a
1345), que se eleva al 76% en el fragmento de 215 nucleótidos comprendido entre el
nucleótido 1372 y el nucleótido 1586 de SEQ ID NO:1; el fragmento comprendido entre
los nucleótidos 1346 y 1371, sin embargo, no aparece entre los de mayor identidad con
25 la secuencia con el número 24 de la familia de patentes de la solicitud internacional
WO 2009158470 y la patente de EE.UU. US8461421.

Sin embargo, tal como se mencionó previamente en la presente memoria, ninguno de los
documentos citados muestra ensayos realizados en los que se demuestre actividad
insecticida de alguna de las proteínas codificadas por dichas secuencias de nucleótidos,
30 asumiéndose que tal actividad existe por responder todas las proteínas a la estructura de
las delta-endotoxinas. No hay tampoco en dichos documentos ninguna sugerencia de que
alguna de esas proteínas codificadas por secuencias con un porcentaje de identidad
considerable con el gen de la presente invención pudiera ser especialmente adecuada
para áfidos. Es más, la secuencia 169 de WO2010099365 pertenece al grupo de genes
35 sintéticos descritos en dicha solicitud de patente que carecen del dominio "cristalino" del

extremo carboxilo que está presente en muchas delta-endotoxinas, y que se cree que está implicado en la formación de cuerpos cristalinos de inclusión dentro de Bt.

La situación es similar en lo que se refiere a otras secuencias de nucleótidos que presentan también un porcentaje de identidad igual o superior al 70% con respecto a la
5 secuencia del gen de la presente invención, representada por SEQ ID NO:1. Así, la CDS (secuencia codificante) con número de acceso KC156704, que se dice que corresponde a una proteína plaguicida de *Bacillus thuringiensis*, cepa ARP242, presenta también un porcentaje de identidad del 87% con respecto a SEQ ID NO:1, pero no parece presentar entre la información disponible a partir de su número de acceso la mención de ninguna
10 publicación o documento de otro tipo en el que se demuestre su actividad plaguicida, no habiendo mención ninguna de posible aplicación a hemípteros, áfidos en particular, o *Myzus persicae* en particular. Igual sucede con las CDS correspondientes a las cepas ARP256 (87% de identidad) o ARP277 (70% de identidad) de Bt. Finalmente, además de fragmentos de la secuenciación del genoma completo de *Bacillus subtilis*, aparecen
15 también porcentajes de identidad altos, pero que no llegan al 75% (74%) con fragmentos reducidos (152 nucleótidos) de secuencias codificantes (CDS) que, hipotéticamente, podrían codificar proteínas Cry a las que se les está asignando una aplicación no plaguicida, sino la de matar células cancerosas.

La comparación realizada utilizando directamente la secuencia de la propia proteína de la
20 presente invención (SEQ ID NO:2) da lugar a que los porcentajes de identidad se reduzcan. Entre las secuencias extraídas de documentos de patentes (Tabla 3), el mayor porcentaje de identidad no llega al 80% (76,20%) y corresponde a la secuencia 87 de la familia de la patente US8318900 y la solicitud de patente internacional WO2010099365, que es precisamente la secuencia de aminoácidos codificada por la secuencia 38 de
25 dichas solicitudes, es decir, la secuencia de nucleótidos con la que SEQ ID NO:1 presenta el mayor porcentaje de identidad. A continuación, sólo la secuencia 84 de la familia de la patente de EE.UU. US8461421 y la solicitud de patente internacional WO2009158470 presenta un porcentaje de identidad superior al 50% (57,91%). El resto de las secuencias extraídas de solicitudes de patente, pertenecientes todas ellas a esta
30 última familia citada, presentan porcentajes de identidad inferiores al 40% con respecto a SEQ ID NO:2.

La comparación con otras proteínas no redundantes no permite localizar proteínas con mayor grado de identidad con la secuencia de la delta-endotoxina de la presente invención. Sólo la proteína similar a Cry41Aa1 (número de acceso en Genbank

AEH76822.1) tiene un porcentaje de identidad cercano al 60%, 57,91%. Además de esta proteína, sólo una secuencia, que corresponde a una proteína hipotética (número de acceso en Genbank EJR94916.1) presenta un porcentaje de identidad mayor al 50% con SEQ ID NO:2.

- 5 Así, ni la búsqueda comparativa por la secuencia del gen ni la búsqueda por la secuencia de la proteína permite localizar secuencias que correspondan a proteínas Cry con actividad conocida contra hemípteros (especialmente contra áfidos) y que presenten un porcentaje de identidad alto (al menos del 50%) con el gen y/o la proteína de la presente invención. En particular, no aparece en ninguna de las Tablas 2, 3 ó 4 la secuencia
10 correspondiente a la proteína Cry o los derivados de la misma con los que se realizan ensayos de actividad insecticida en la familia de patentes de la solicitud internacional WO2010099365 y la patente de EE.UU. US8318900B2.

La delta-endotoxina de la presente invención es bastante diferente de otras proteínas Cry activas contra áfidos descritas hasta el momento. Tal como se muestra en la Tabla 5 que
15 aparece más adelante en la presente memoria, la proteína Cry con actividad conocida contra áfidos que presenta mayor porcentaje de identidad con la proteína Cry de la presente invención es la proteína Cry3Aa, con la que sólo comparte un 26,3% de identidad. Con respecto a otras proteínas para las que se haya ensayado su posible actividad contra áfidos (Cry2Aa, Cry4Aa, Cry11Aa e, incluso, la proteína representada por
20 la secuencia con número de identificación 205 en la patente US8318900B2 y de la solicitud de patente internacional WO2010099365), la identidad no llega al 20%.

Y, sin embargo, sorprendentemente, los ensayos experimentales descritos más adelante en Ejemplos de la presente solicitud demuestran que la nueva proteína Cry de la presente invención es la primera proteína Cry de Bt con actividad insecticida significativa
25 contra áfidos en general y, en particular, la primera proteína Cry de Bt para la que se divulga que posea actividad insecticida contra el pulgón verde, *Myzus persicae*, por lo que, en consecuencia, es un objeto de la presente invención, junto con el gen que la codifica, así como el uso como insecticida de las composiciones que contienen el gen y/o la proteína, incluida la generación de plantas transgénicas que expresen dicho gen.

30 Es más, los resultados experimentales demuestran que esta proteína presenta una elevada actividad insecticida siendo su CL₅₀ de tan solo 32,7 µg/ml. Este valor es mucho menor que los valores obtenidos con otras proteínas Cry previamente evaluadas en otras especies de áfidos relacionadas con *M. persicae* (Tabla 1). Es también significativa la diferencia en orden de magnitud con las cantidades de otras proteínas Cry utilizadas en

ensayos como los que se describen en la solicitud de patente internacional WO9516778 donde algunas proteínas análogas, como la mezcla de proteínas CryI, llega a utilizarse a una concentración próxima a 400 mg/ml, y donde otras proteínas, como CryIII B2, utilizada a una concentración casi un orden de magnitud superior (150 µg/ml), necesitan hasta 7 días para dar lugar a una mortalidad cercana al 100% en ensayos dirigidos a áfidos, en los que *Myzus persicae* no se menciona. Incluso extendiendo el ámbito de comparación a las actividades conocidas sobre otros hemípteros, el valor puede considerarse elevado, como puede comprobarse a partir del Ejemplo 18 de la solicitud WO2010099365 y de la patente US8318900B2, realizado con derivados no naturales de la proteína representada por la secuencia 205 de dichos documentos, donde la concentración de 50 µg/ml da lugar a un 80% de mortalidad del hemíptero *Lygus hesperus*, de la familia *Miridae*, siendo necesario un extracto concentrado obtenido de *E. coli* para que uno de los derivados de la misma proteína dé lugar al 100% de mortalidad en el 25-50% de los recipientes cuando el ensayo se realiza con un áfido, *Aphis glycines*.

En general, como se comentó previamente en la presente solicitud, los áfidos no resultan muy susceptibles a las toxinas Cry de Bt (Chougule and Bonning, 2012). Este hecho, junto a la capacidad de los áfidos para generar resistencias a los insecticidas químicos de síntesis, incrementa la importancia del descubrimiento de nuevas proteínas más tóxicas para implementar métodos alternativos de control basados en toxinas Bt. La nueva proteína Cry de la presente invención se presenta como tal alternativa, pudiendo ser adaptada al desarrollo de métodos de control basados en nuevos formulados bioinsecticidas y la construcción de plantas transgénicas resistentes al pulgón verde y otras especies de áfidos susceptibles.

La secuencia codificante de esta nueva proteína puede utilizarse para transformar huéspedes heterólogos bacterianos o células vegetales mediante técnicas bien conocidas y rutinarias en ingeniería genética, bien para producir nuevas copias de la proteína o como paso intermedio para la generación de plantas transgénicas que la expresen. Para ello, un paso previo suele ser su inserción en un "vector", una molécula de ADN que posibilita que se lleve a cabo el objetivo que se pretende conseguir, bien la expresión del gen (vectores de expresión) o bien la transferencia de moléculas o construcciones de ADN entre distintas células hospedadoras, como puede ser la transformación de células de plantas, a menudo incluyendo la posterior obtención de una planta transgénica (los llamados vectores de transformación, que a menudo son también vectores de expresión, o incluyen lo que se llaman "casetes de expresión"). Tanto en uno como en otro casos, es frecuente que un paso previo sea la generación de "casetes de expresión", en las que el

gen a expresar está operativamente unido a un promotor que permite su expresión en un sistema de interés y que a menudo contienen también alguna secuencia regulatoria adicional o de utilidad para la expresión, como una región no traducida de 3' o, especialmente cuando se desea la expresión en células eucarióticas, una secuencia señal o una secuencia líder que facilite el transporte del péptido que se está formando, o una vez traducido, a un orgánulo celular determinado, tales como cloroplastos, retículo endoplásmico, aparato de Golgi...; también es habitual que el "casete de expresión" se forme al insertar la secuencia codificante a expresar en el vector en un lugar preciso, de manera que quede unida a los elementos deseados.

- 5
- 10 La elección del promotor depende del sistema en el que se desee que se exprese el gen al que va a quedar operativamente unido. Así, por ejemplo, cuando la proteína va a producirse, por ejemplo, en un microorganismo, tal como una bacteria, será necesario elegir un promotor que permita la expresión en dicha bacteria. Una de las bacterias más utilizadas para la expresión de proteínas es *Escherichia coli*, para la cual hay muchos
- 15 promotores y vectores de expresión que los contienen bien conocidos por los expertos en la materia. En el caso de *Escherichia coli*, es habitual insertar los genes a expresar bajo el control de operón lac, como en el Ejemplo 2 de la presente invención, lo cual permite controlar la expresión de la proteína, induciéndola mediante la adición de IPTG (isopropil- β -D-1-tiogalactopiranosido). Son muchos los vectores de expresión conocidos y
- 20 disponibles comercialmente que permiten la expresión en *E. coli*, de los cuales son muy utilizados los plásmidos, pero también son habituales y conocidos algunos virus recombinantes preparados a partir de las formas naturales que infectan dicha bacteria, como el bacteriófago lambda. Los vectores de expresión tipo plásmido son fáciles de localizar y de elegir a partir no sólo de catálogos comerciales, sino de bases de datos
- 25 dedicadas a este tipo de vectores como la de Addgene (<http://www.addgene.org>), donde puede consultarse, por ejemplo, la estructura del plásmido pET-28b(+) utilizado en los ejemplos de la presente invención.

Otros huéspedes microbianos disponibles para la transformación con vectores de expresión, de especial interés en el presente caso, podrían ser la cepa Bt 4Q7

30 acristalófera (no productoras de otras endotoxinas) disponible en el Bacillus Genetic Stock Center, *Bacillus megaterium* (Mobictec, Germany), *Pseudomonas* spp. o levaduras. En estos microorganismos, la proteína insecticida producida permanece encapsulada en el soma bacteriano de tal manera que el mismo microorganismo se convertiría en la materia activa del biopesticida, fácilmente recuperable por centrifugación. Esta

35 microencapsulación, protegería además a la proteína de factores ambientales (ej. UV)

alargando su vida útil en el medioambiente. La opción de recubrir dichas proteínas con sustancias (nanopartículas) que protejan de manera similar a dicha toxina resulta también una alternativa atractiva.

5 Este nuevo gen *cry*, puede también optimizarse para su expresión en células vegetales en la construcción de plantas transgénicas resistentes al insecto, mediante la utilización de técnicas ya descritas y bien conocidas de ingeniería genética.

10 Los métodos de transformación de plantas de la invención implican la introducción de una construcción nucleotídica, incluyendo un gen a transferir, en una planta. Por "introducción" se entiende proporcionar a la planta la construcción de nucleótidos de tal manera que la construcción acceda al interior de una célula de la planta. Los métodos de la presente invención no requieren que se utilice un método particular para la introducción de una construcción de nucleótidos en una planta. Los métodos para introducir construcciones de nucleótidos en plantas son conocidos en la técnica, incluyendo, pero sin limitarse a, métodos de transformación estable, métodos de transformación transitoria, 15 y métodos mediados por virus.

Por "planta" se entiende plantas completas, órganos de plantas (por ejemplo, hojas, tallos, raíces, etc), semillas, células de plantas, propágulos, embriones y la progenie de los mismos. Las células de plantas (o células vegetales) pueden ser diferenciados o indiferenciados (por ejemplo, callos, las células en cultivo en suspensión, protoplastos, 20 células de las hojas, células de la raíz, células del floema, polen).

Los términos "plantas transgénicas" o "plantas transformadas" o plantas células o tejidos "transformados de forma estable" se refieren a plantas que han incorporado o integrado secuencias de ácidos nucleicos o fragmentos de ADN exógenos en una o más células vegetales. Estas secuencias de ácidos nucleicos incluyen las que son exógenos, o no 25 están presentes en la célula de la planta no transformada, así como aquellas que pueden ser endógenas, o que están presentes en la célula de la planta no transformada. Con el término "heteróloga" generalmente se hace referencia a las secuencias de ácido nucleico que no son endógenas para la célula o parte del genoma nativo en el que están presentes, y que se han añadido a la célula por infección, transfección, microinyección, electroporación , microproyección, u otros similares. 30

Las plantas transgénicas de la presente invención serán aquellas que expresen el gen de la presente invención o una secuencia codificante de una delta-endotoxina que presente al menos un 90% de identidad con la secuencia del gen de la presente invención (SEQ ID

NO:1). La planta transgénica puede comprender además uno o más genes adicionales para la resistencia a los insectos, y/o cualquier gen adicional que imparta un rasgo agronómico de interés.

La transformación de células vegetales se puede lograr mediante diversas técnicas bien conocidas por el experto en la materia. Normalmente, el gen que se quiere expresar (en este caso, el gen de la presente invención, o secuencias codificantes que presenten al menos un 90% de identidad con el mismo), estará formando parte de una construcción (un casete de expresión, como se describió anteriormente) en la que el gen estará unido operativamente a un promotor que controlará su transcripción, así como a una zona no traducida de 3' que permitirá la terminación de la transcripción y la poliadenilación del transcrito. Entre las posibles regiones de terminación pueden ser adecuadas, por ejemplo, las que están presentes en el plásmido Ti de *Agrobacterium tumefaciens*, tales como las regiones de terminación de la octopina sintasa y la nopalina sintasa. Si se desea, puede incluirse también un péptido señal que facilite la transferencia del péptido que se está formando al retículo endoplásmico. También puede ser interesante diseñar el casete de expresión de la planta para que contenga al menos un intrón, de tal manera que se requiera el procesamiento del intrón del ARNm para su expresión.

Normalmente, este "casete de expresión en plantas" se inserta o está inserto en un "vector de transformación de plantas". Este vector de transformación de plantas puede estar compuesto de uno o más vectores de ADN necesarios para lograr la transformación de la planta. Por ejemplo, es una práctica común en la técnica utilizar vectores de transformación de plantas que se componen de más de un segmento de ADN contiguo. Estos vectores se denominan a menudo "vectores binarios". Los vectores binarios, así como los vectores con plásmidos auxiliares, son los más utilizados para la transformación mediada por *Agrobacterium*, cuando el tamaño y la complejidad de los segmentos de ADN necesarios para lograr la transformación eficiente es bastante grande. Los vectores binarios contienen típicamente un vector plasmídico que contiene las secuencias que actúan en cis necesarias para la transferencia de T-ADN (tal como el borde izquierdo y el borde derecho), un marcador de selección que está diseñado para ser capaz de expresarse en una célula vegetal, y un "gen de interés" (un gen diseñado para ser capaz de expresarse en una célula de una planta de la cual se desea la generación de plantas transgénicas que, en el presente caso, sería el gen de la presente invención o una secuencia codificante que guarde al menos un 90% de identidad con el mismo). También están presentes en este vector plasmídico secuencias requeridas para la replicación en bacterias. Las secuencias que actúan en cis están dispuestas de tal manera que permitan

la transferencia eficiente en células vegetales y la expresión en las mismas. Por ejemplo, el gen marcador de selección y el gen de la presente invención se encuentran entre los bordes izquierdo y derecho. A menudo un segundo vector plasmídico contiene los factores que actúan en trans que median la transferencia del T-ADN de *Agrobacterium* a las células vegetales. Este plásmido contiene a menudo las funciones de virulencia (genes Vir) que permiten la infección de las células vegetales por *Agrobacterium*, y la transferencia de ADN mediante escisión de las secuencias de los bordes y la transferencia de ADN mediada por vir, como se entiende en la técnica (Hellens and Mullineaux, 2000). Para la transformación de plantas pueden utilizarse varios tipos de cepas de *Agrobacterium* (por ejemplo, LBA4404, GV3101, EHA101, EHA105, etc). El segundo vector plasmídico no es necesario para la transformación de las plantas por otros métodos tales como microproyección, microinyección, electroporación, polietilenglicol, etc.

En general, los métodos de transformación de plantas involucran la transferencia de ADN heterólogo a las células diana de la planta (por ejemplo, embriones inmaduros o maduros, cultivos en suspensión, callo indiferenciado, protoplastos, etc), seguido por la aplicación de un nivel umbral máximo para la selección apropiada (dependiendo del gen marcador seleccionable) para recuperar las células de plantas transformadas a partir de un grupo de una masa celular sin transformar. Los explantes se transfieren normalmente a un suministro fresco del mismo medio y se cultivan de forma rutinaria. Posteriormente, las células transformadas se diferencian en brotes después de colocarlas en medio de regeneración suplementado con un nivel de umbral máximo de agente de selección. Los brotes son luego transferidos a un medio selectivo de enraizamiento para la recuperación de brotes con raíces o plántulas enraizadas. La plántula transgénica crece entonces dando lugar a una planta madura y produce semillas fértiles. Los explantes se transfieren normalmente a un suministro fresco del mismo medio y se cultivan de forma rutinaria. Una descripción general de las técnicas y métodos para generar plantas transgénicas se encuentran en publicaciones como la de Ayres y Park (Ayres and Park, 1994) y Bommineni y Jauhar (1997). Dado que el material transformado contiene muchas células, cualquier muestra de dicho material transformado contiene células transformadas y no transformadas. La capacidad de matar células no transformadas y permitir que las células transformadas proliferen da lugar a que sean convenientes los cultivos de plantas transformadas. A menudo, la capacidad de eliminar las células no transformadas es una limitación para la rápida recuperación de las células de plantas transformadas y la generación exitosa de plantas transgénicas.

Los protocolos de transformación así como los protocolos para introducir secuencias de nucleótidos en plantas pueden variar dependiendo del tipo de planta o célula vegetal que se quiera transformar, es decir, monocotiledónea o dicotiledónea. Como se ha mencionado previamente, la generación de plantas transgénicas puede ser realizada por diversos métodos, incluyendo, pero sin limitarse a, microinyección, electroporación, transferencia directa de genes, introducción de ADN heterólogo por medio de *Agrobacterium* en células de plantas (transformación mediada por *Agrobacterium*), bombardeo de células vegetales con ADN adherido a partículas, la aceleración de partículas balísticas, transformación mediante un haz de aerosol, la transformación con Lec1, y diversos otros métodos no mediados por partículas.

Tras la integración del ADN foráneo heterólogo en células de la planta, se aplica entonces un nivel umbral máximo de selección apropiada en el medio para eliminar las células no transformadas y separar y hacer proliferar las células hipotéticamente transformadas que sobreviven a este tratamiento de selección transfiriéndolas periódicamente a un medio fresco, mediante pases continuos y manteniendo las apropiadas condiciones de selección. Se han desarrollado numerosos marcadores para su uso con células vegetales, tales como la resistencia a cloranfenicol, al aminoglicósido G418, resistencia a higromicina, o similares.

Se pueden utilizar métodos moleculares y bioquímicos para confirmar la presencia del gen heterólogo de interés integrado en el genoma de la planta transgénica, gen de interés que, en el caso de las plantas transgénicas de la presente invención, es el gen de la presente invención o una secuencia codificante con un porcentaje de identidad de al menos el 90% con el mismo.

Las células que han sido transformadas pueden ser cultivadas para dar lugar a plantas de forma convencional. Estas plantas pueden ser cultivadas a continuación, y, o bien ser polinizadas con la misma cepa transformada o con cepas diferentes, y el híbrido resultante tiene expresión constitutiva de la característica fenotípica deseada, en el presente caso, la expresión de la proteína Cry de la presente invención. Dos o más generaciones se pueden cultivar para asegurar que la expresión de la característica fenotípica deseada se mantiene y se hereda de manera estable y luego cosechar las semillas para asegurar que se ha logrado la expresión de la característica fenotípica deseada. De esta manera, la presente invención proporciona semillas transformadas (a las que también se hace referencia como "semillas transgénicas") que tiene una construcción nucleotídica de la presente invención, (por ejemplo, un casete de expresión

en el que el gen de la invención está ligado, de manera estable, a un promotor que permite su expresión en plantas y a una secuencia adecuada de terminación, incorporado de manera estable en su genoma).

Tras la introducción de ADN foráneo heterólogo en células de plantas, la transformación o integración del gen heterólogo en el genoma de la planta se puede confirmar por diversos métodos tales como el análisis de ácidos nucleicos, proteínas y metabolitos asociados con el gen integrado. En este caso en particular, se puede realizar el análisis de la presencia del gen integrado, de la generación de su ARNm o de la presencia de la proteína que se quiere expresar, la proteína de la presente invención o una proteína que presenta un alto grado de identidad con la misma.

El análisis por PCR es un método rápido para detectar en células, tejidos o brotes transformados, la presencia del gen incorporado en una etapa anterior antes de trasplantarlo en el suelo, mediante métodos rutinarios para los expertos en la materia (véase, por ejemplo, Sambrook y Russell, 2001). La PCR se puede llevar a cabo utilizando los oligonucleótidos cebadores específicos que permitan amplificar el gen de interés desde el codón de inicio (ATG) hasta el codón de parada de la traducción (TTA) (oligonucleótido cebador directo 5'-ATGAACCCAAAATTATAACAACAATG-3', (SEQ ID NO:3) oligonucleótido cebador inverso 5' CAAATCAAAAATTCAAAGTGAATAAGTTA3' (SEQ ID NO:4)). Mediante amplificación por PCR se podría además obtener el amplicón, purificarlo y marcarlo ya sea mediante métodos radioactivos (³²P) o no radioactivos (digoxigenina) y utilizarlo como sonda específica.

La transformación de plantas puede ser confirmada por análisis mediante transferencia tipo Southern de ADN genómico (Sambrook y Russell, 2001). En general, el ADN total se extrae a partir del transformante, se digiere con enzimas de restricción apropiadas, se fracciona en un gel de agarosa y se transfiere a una membrana de nitrocelulosa o de nylon. A continuación, se realiza un ensayo con la membrana con el ADN transferido con, por ejemplo, un fragmento del ADN diana (un fragmento del gen de la presente invención, en este caso, o una secuencia que tiene al menos un 90% de identidad con el mismo) para confirmar la integración del gen introducido en el genoma de la planta.

En el análisis de transferencia tipo Northern, se aísla ARN a partir de tejidos específicos del transformante, se fraccionan en un gel de agarosa-formaldehído, y se transfiere a un filtro de nylon de acuerdo con procedimientos estándar que se utilizan rutinariamente por los expertos en la técnica (Sambrook y Russell, 2001). La expresión del ARN codificado por el gen de la presente invención o por una secuencia codificante que tiene al menos

un 90% de identidad con el mismo se prueba mediante la hibridación del filtro con una sonda radioactiva derivada de la secuencia del gen de la endotoxina de la invención, por métodos conocidos en la técnica (Sambrook y Russell, 2001).

5 También se puede confirmar la presencia de la proteína codificada por el gen de la delta-endotoxina de la presente invención mediante transferencia tipo Western, ensayos bioquímicos o similares llevados a cabo en las plantas transgénicas por procedimientos estándar (Sambrook y Russell, 2001, *supra*), utilizando anticuerpos que se unen a uno o más epítomos presentes en la proteína de la delta-endotoxina de la presente invención. Para ello, pueden utilizarse, por ejemplo, anticuerpos dirigidos a otras
10 proteínas Cry con las que la proteína Cry de la presente invención comparta epítomos. Otra alternativa es la preparación de anticuerpos específicos contra la nueva proteína Cry de la presente invención y, de una manera similar, utilizar anticuerpos Anti-6xHis_Tag (Sigma-Aldrich) contra la cola de 6 histidinas fusionadas al extremo carboxilo de la proteína recombinante CyBM1 y que están codificadas en el vector pET-28b(+). La
15 generación de anticuerpos frente a una proteína cualquiera es una técnica habitual, conocida por cualquier experto en la técnica, que puede considerarse rutinaria.

Se tiene preferencia por los anticuerpos monoclonales, que son anticuerpos homogéneos producidos por una célula híbrida producto de la fusión de un clon de linfocitos B descendiente de una sola y única célula madre y una célula plasmática tumoral, que
20 garantiza su perpetuidad (Köhler & Milstein, 1975). Con esta fusión de dos células, una programada para producir un anticuerpo específico pero que no se multiplica indefinidamente (linfocito) y otra inmortal con gran capacidad de crecimiento pero que no produce inmunoglobulina (célula de mieloma), se combina la información genética necesaria para la síntesis del anticuerpo deseado y una capacidad de síntesis proteica,
25 permitiendo su multiplicación indefinida tanto *in vitro* como *in vivo*. Para la producción de anticuerpos monoclonales frente a la proteína Cry de la presente invención, los presentes autores de la presente invención tienen preferencia por la metodología descrita por Iglesias *et al.* (1997). Brevemente: se fusionan esplenocitos obtenidos de ratones BALB/c hembras, previamente inyectadas intraperitonealmente con la forma completa de
30 la proteína Cry de la presente invención, con células de mieloma murino X63-Ag8.653, en presencia de 50% de polietilenglicol (PEG) 4000 (Boehringer Mannheim, Barcelona, España). Los cultivos de hibridomas se pueden mantener según lo descrito por Estévez *et al.* (1994). Posteriormente, se obtienen los anticuerpos monoclonales anti-Cry a partir del líquido ascítico de ratones inyectados, vía intraperitoneal, con las células del
35 hibridoma seleccionado (2×10^6 células/ratón, por ejemplo). La purificación parcial de los

anticuerpos se puede realizar mediante precipitación de la ascitis con sulfato amónico saturado (SAS). En cualquier caso, actualmente en España son muchas las compañías biomédicas especializadas en la producción de anticuerpos monoclonales a la carta, a las que se puede encargar su producción.

- 5 En el presente caso, la presencia del transgén de la delta-endotoxina de la invención se puede detectar también mediante pruebas de su actividad plaguicida en plantas fértiles. Las plantas que muestran actividad óptima se seleccionan para su posterior reproducción. Hay métodos disponibles en la técnica para ensayar la actividad de plagas como, por ejemplo, mezclar la proteína y utilizarla en ensayos de alimentación, como en
10 los Ejemplos que se muestran más adelante en la presente solicitud.

La utilización de cualquiera de las metodologías anteriormente descritas dan lugar a plantas transgénicas que expresan la delta-endotoxina de la presente invención, por lo que dichas plantas transgénicas tendrán actividad insecticida contra, al menos, un áfido, preferiblemente *Myzus persicae*, de manera que el uso de dichas plantas transgénicas
15 para combatir plagas de áfidos y/o conferir a las plantas resistencia a los mismos, en particular de *Myzus persicae*, es también un objeto de la presente invención. En principio, la manera en que se generan las células de plantas transgénicas no es crítico para los fines de la presente invención. Sin embargo, puesto que el objetivo principal de la presente invención es combatir las plagas de hemípteros, preferiblemente áfidos, de
20 entre ellos *Myzus persicae* en particular, es una realización especialmente interesante de la presente invención aquella en la que las plantas transformadas (transgénicas) selectivamente expresan la delta-endotoxina de la presente invención en el tejido vascular de la planta, en particular en el tejido del floema, como se divulga en la solicitud WO95/16778, para facilitar la ingestión de la delta-endotoxina por parte de los áfidos a
25 combatir. Esto puede conseguirse eligiendo como promotor operativamente unido al gen de la presente invención, o a una secuencia codificante con un porcentaje de identidad de al menos el 90% con el mismo, un promotor que facilite la expresión en dichos tejidos. Para ello, son promotores adecuados los mismos mencionados en dicha solicitud de patente internacional, tales como el promotor del virus del mosaico de la coliflor 4xB2+A
30 CaMV35S.

La presente invención se puede utilizar para la transformación de cualquier tipo de planta. Las plantas por las que se tiene especial preferencia son aquellas que sufren daños causados por áfidos, en particular por *Myzus persicae*, o en las que se ha comprobado que dicho pulgón facilita la transmisión de virus. Así, entre otras, se tiene preferencia por

plantas transgénicas de melocotonero (*Prunus persica*), albaricoquero, ciruelo y otras plantas del género *Prunus*, cítricos, crucíferas, solanáceas, gramíneas e, incluso, plantas ornamentales.

- Alternativamente, las plantas pueden protegerse mediante la aplicación de una composición que contiene la delta-endotoxina de la invención o un sistema capaz de expresarla. También pueden utilizarse composiciones que comprenden la cepa de *Bacillus* que contiene en su genoma la secuencia de nucleótidos de la cual se ha aislado el gen de la presente invención, *Bacillus thuringiensis* Bt H1.5, u otros microorganismos recombinantes en los que se exprese la delta-endotoxina de la invención.
- 10 Los ingredientes activos (la delta-endotoxina de la presente invención o un microorganismo que la exprese) de la presente invención se aplican normalmente en forma de composiciones (las composiciones plaguicidas de la presente invención) y pueden aplicarse al área de cultivo o planta a tratar, simultáneamente o de forma consecutiva, con otros compuestos. Estos compuestos pueden ser fertilizantes, herbicidas, crioprotectores, agentes tensioactivos, detergentes, jabones, aceites, plaguicidas latentes, polímeros, o formulaciones de vehículos biodegradables que permiten la dosificación a largo plazo en un área de destino después de una sola aplicación de la formulación. También pueden ser herbicidas selectivos, insecticidas químicos, virucidas, microbicidas, amibicidas, pesticidas, fungicidas, bactericidas, nematocidas, molusquicidas o mezclas de varias de estas preparaciones, administradas, si se desea, junto con otros vehículos agrícolamente aceptables, tensioactivos o adyuvantes de la aplicación que se emplean habitualmente en la técnica de la formulación de plaguicidas. Los vehículos y adyuvantes adecuados pueden ser sólidos o líquidos y corresponderse con las sustancias empleadas normalmente en la tecnología de la formulación, por ejemplo, sustancias minerales naturales, disolventes, dispersantes, agentes humectantes, aglutinantes o fertilizantes. Del mismo modo las formulaciones se pueden preparar en forma de "cebos comestibles" o formar "trampas" para plagas para permitir la alimentación o la ingestión por una plaga diana de la formulación plaguicida; un ejemplo de tales dispositivos puede encontrarse en la solicitud WO9516778.
- 20
- 25
- 30 Los métodos de aplicación de una composición de la presente invención, que comprende al menos la delta-endotoxina de la presente invención o un sistema de expresión de la misma, tal como células de *Bacillus thuringiensis* Bt H1.5 o de un microorganismo recombinante que comprende el gen de la invención operativamente unido a un promotor que permite la expresión en dicho microorganismo, incluyen la aplicación sobre las hojas,

recubrimiento de semillas y aplicación en el suelo. La composición se puede formular en forma de polvo, gránulo, aerosol, emulsión, coloide, solución, o similares, y se puede preparar por medios convencionales tales como la desecación, liofilización, homogeneización, extracción, filtración, centrifugación, sedimentación, o concentración
5 de un cultivo de células que comprenden el polipéptido correspondiente a la delta-endotoxina de la presente invención.

Las formulaciones también pueden variar con respecto a las condiciones climáticas, las consideraciones ambientales, la frecuencia deseada de aplicación y/o la severidad de la infestación.

10 Las composiciones descritas pueden efectuarse mediante la formulación de cualquier célula bacteriana de las anteriormente citadas, (que, en el caso de *Bacillus thuringiensis*, puede ser, por ejemplo, una suspensión de esporas o de formas cristalinas), o la formulación de la proteína de la presente invención, previamente producida y aislada, con el vehículo agrícolamente aceptable deseado.

15 Un ejemplo de composición que comprende la delta-endotoxina de la presente invención puede encontrarse en el Ejemplo 3 de la presente solicitud, donde la proteína soluble se diluyó en una disolución acuosa que contenía sacarosa, concretamente sacarosa al 20% p/v, que era la dieta artificial preparada para los áfidos. Otros ejemplos de formulación pueden encontrarse, por ejemplo, en la solicitud internacional WO9516778, donde las
20 delta-endotoxinas están en forma cristalina y se preparan en forma de suspensión acuosa.

En general, las composiciones se pueden formular antes de la administración en un medio apropiado, y se pueden preparar en la forma deseada por medios conocidos tales como liofilización, secado con congelación posterior, o desecación, o en un vehículo
25 acuoso, medio o diluyente adecuado, tal como solución salina u otro tampón. Las composiciones formuladas pueden estar en la forma de un polvo o material granular, o una suspensión en aceite (vegetal o mineral), o agua o emulsiones de aceite / agua, o como un polvo humectable, o en combinación con cualquier otro vehículo adecuado para aplicación agrícola. El término "vehículo agrícolamente aceptable" cubre todos los
30 adyuvantes, componentes inertes, dispersantes, tensioactivos, adhesivos, aglutinantes, etc, que se utilizan habitualmente en la tecnología de formulación de plaguicidas, que son bien conocidos para los expertos en formulación plaguicida. Las formulaciones se pueden mezclar con uno o más adyuvantes sólidos o líquidos y prepararse por diversos medios, por ejemplo, mezcla hasta homogeneización, mezcla y/o molienda de la composición

plaguicida con adyuvantes adecuados usando técnicas de formulación convencionales. Formulaciones adecuadas y los métodos de aplicación se describen, por ejemplo, en la Patente de EE.UU. Nº 6468523, incorporada a la presente solicitud por referencia.

5 Otra posible alternativa es que el gen de la presente invención se clone para su expresión en la cepa Bt 4Q7, o en microorganismos como *Bacillus megaterium*, *Pseudomonas* spp., o levaduras, en los que la delta-endotoxina producida permanece encapsulada en el soma bacteriano, de tal manera que el mismo microorganismo se convertiría en la materia activa del biopesticida, fácilmente recuperable por centrifugación. Esta microencapsulación, protegería además a la proteína de factores ambientales (ej. UV) alargando su vida útil en el medioambiente. La opción de recubrir dichas proteínas con sustancias (nanopartículas) que protejan de manera similar a la delta-endotoxina es también una posible alternativa.

15 La delta-endotoxina microencapsulada se puede utilizar en aplicaciones de pulverización, bien de forma independiente, o en rotaciones con insecticidas a base de Bt que exprese otras posibles delta-endotoxinas de interés.

Las plantas también pueden ser tratadas con una o más composiciones químicas, incluyendo uno o más herbicidas, insecticidas, o funguicidas.

20 Por todo ello, la presente invención se refiere a una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica una delta-endotoxina, caracterizada por que la secuencia de nucleótidos que codifica una delta-endotoxina está representada por SEQ ID NO:1 o presenta al menos un 90% de identidad con la misma. En posibles realizaciones, el porcentaje de identidad puede seleccionarse, por ejemplo, del grupo de 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5%, 99,6%, 99,7%, 99,8%, 99,9%, porcentajes de identidad que se consideran que dan lugar a una delta-endotoxina suficientemente idéntica a la delta-endotoxina cuyo aislamiento y producción se describe más adelante en Ejemplos de la presente invención. Lógicamente, otra posible realización es aquella en la que la identidad es del 100%, realización que correspondería a la alternativa en la que la secuencia de nucleótidos que codifica una delta-endotoxina es exactamente la representada por SEQ ID NO:1, siendo así una posible realización de este aspecto de la invención que la secuencia de nucleótidos de la molécula de ácido nucleico esté representada exactamente por SEQ ID NO:1.

La determinación del porcentaje de identidad entre dos secuencias se puede llevar a

cabo con los recursos informáticos habituales, basados en algoritmos matemáticos, como pueden ser los programas BLASTN y BLASTP utilizados en los Ejemplos posteriores de la presente solicitud, utilizando, por ejemplo, los parámetros que aparecen por defecto, así como con otras aplicaciones alternativas conocidas por los expertos en la materia,
5 como Gapped Blast y PSI-Blast (Altschul et al., 1997), u otros.

Tal como se comenta más adelante, la proteína Cry de la presente invención presenta un dominio tipo ricina b, que no está presente en otras proteínas Cry con actividad contra áfidos. Por ello, se prefiere que se mantenga que la secuencia de nucleótidos que codifica una delta-endotoxina comprende un fragmento de secuencia que codifica un
10 dominio tipo ricina b que presenta al menos un 90% de identidad con la secuencia codificante de los aminoácidos 652-799 de SEQ ID NO:2.

Son realizaciones preferidas de la presente invención, compatibles con las anteriores, aquellas en las que la secuencia de nucleótidos que codifica una delta-endotoxina tiene actividad contra al menos una especie de insecto hemíptero. Se prefiere especialmente
15 que presente actividad contra al menos una especie de áfido y, muy especialmente, que tenga actividad contra el áfido *Myzus persicae*.

Otro aspecto de la invención lo constituyen las "construcciones" o "casetes de expresión" que posibilitan que el gen se exprese en un sistema de expresión u hospedador de interés, así como los vectores que contienen dichas construcciones y, que facilitan la
20 introducción del gen en dichos sistemas y/o su expresión en los mismos.

Por ello, la invención se refiere también a una molécula de ácido nucleico aislada como cualquiera de las definidas anteriormente como parte de la presente invención, en la que la secuencia de nucleótidos que codifica la delta-endotoxina de la invención está operativamente unida a un promotor. Según donde se desee expresar el gen (la
25 secuencia de nucleótidos que codifica la delta-endotoxina de la invención), se tendrá preferencia por distintos posibles promotores, que pueden ser promotores capaces de dirigir la expresión de la secuencia en una célula de planta, en una bacteria (preferiblemente seleccionada del grupo de *Escherichia coli*, *Bacillus thuringiensis*, *Pseudomonas spp.*) o en una levadura como *Saccharomyces cerevisiae*. Son posibles realizaciones de este
30 aspecto de la invención aquellas en las que la secuencia codificante de la delta-endotoxina está operativamente unida a alguna secuencia de control adicional, tal como una región no traducida de 3'.

Otro aspecto de la invención son los vectores que comprenden la secuencia codificante

de la delta-endotoxina de la presente invención, preferiblemente ligada a otros elementos de manera que formen las "construcciones" o "casetes de expresión" previamente definidos. Así, otro aspecto de la invención es un vector que comprende una secuencia de nucleótidos que codifica una delta-endotoxina, caracterizada por que la secuencia de nucleótidos que codifica una delta-endotoxina está representada por SEQ ID NO:1 o presenta al menos un 90% de identidad con la misma, o una molécula de ácido nucleico en la que esté comprendida dicha secuencia de nucleótidos, es decir, un vector en el que está inserta la molécula de ácido nucleico que constituye el primer aspecto de la presente invención. El vector se elige según el sistema donde se desea la expresión o que se desea transformar con él. En una realización preferida, el vector es un plásmido, lo que permite obtener un número elevado de copias en bacterias adecuadas transformadas con él y, en caso de que el promotor sea adecuado, permite también la expresión de la delta-endotoxina de la presente invención en la bacteria, siendo posible la producción y purificación de la delta-endotoxina a partir de cultivos bacterianos. También está comprendida dentro del alcance de la invención la realización en la que el plásmido es un plásmido que puede replicarse (y expresarse) en levaduras, como es el caso del llamado plásmido de 2 micras (2 μ), de *Saccharomyces cerevisiae*. Una realización preferida es aquella en la que el plásmido permite la expresión de la delta-endotoxina de la presente invención en una bacteria que se elige entre *Escherichia coli*, *Pseudomonas spp.*, *Bacillus megaterium* o *Bacillus thuringiensis*; en este último caso, se tiene especial preferencia por la cepa Bt 4Q7.

El vector puede ser también un vector de transformación de plantas, realización compatible, entre otras, con la alternativa de que el vector sea un plásmido. Se prefiere especialmente que dicho vector de transformación esté construido de la manera que la secuencia de nucleótidos que codifica la delta-endotoxina representada por SEQ ID NO:1 o la secuencia codificante que presenta al menos un 90% de identidad con la misma está operativamente unida a un promotor capaz de dirigir la expresión de dicha secuencia de nucleótidos en una célula de planta, como puede ser el promotor del virus del mosaico de la coliflor (promotor de CaMV) o la variante 4xB2+A CaMV35S; dentro de esta realización, se prefiere particularmente que la secuencia codificante esté también operativamente unida a otras secuencias de control o regulatorias que faciliten o posibiliten la expresión en plantas, como secuencias de terminación o, incluso, secuencias líder, aunque la presencia de secuencias regulatorias es igualmente compatible con cualquier vector, incluidos los de expresión en bacterias. Con especial preferencia, la estructura del vector de transformación deberá ser tal que

toda la construcción o "casete de expresión" (promotor, secuencia codificante y secuencias regulatorias o de control) pueda resultar transferido como tal conjunto estructural al genoma de la planta. Es también una realización preferida, compatible con todas las anteriores, aquella en la que la secuencia codificante de la delta-endotoxina de la presente invención está unida a la secuencia codificante de un segundo polipéptido, en fase con ella, y todo el conjunto bajo el control del mismo promotor, de manera que la expresión de la sucesión de secuencias codificantes dé lugar a una proteína de fusión.

Otro aspecto más de la invención lo constituyen las células hospedadoras que comprendan al menos un vector de la presente invención. Tal como se ha discutido previamente, entre los posibles hospedadores preferidos están las células de bacterias, muy especialmente las del grupo de *Escherichia coli*, *Pseudomonas spp.*, *Bacillus megaterium* o *Bacillus thuringiensis*, especie esta última en la que se tiene especial preferencia por la cepa Bt 4Q7. La célula puede ser también una célula de planta. También es posible, entre otras realizaciones, que la célula sea una célula de levadura, particularmente de *Saccharomyces cerevisiae*.

Un aspecto más de la presente invención son las plantas transgénicas que comprenden células transformadas con vectores de la presente invención. Por tanto, es un aspecto de la presente invención una planta transgénica que comprende al menos una célula hospedadora que comprende al menos un vector de la presente invención. Por planta, como ya se mencionó anteriormente, puede entenderse una planta completa o cualquier parte de la misma, así como las formas previas en su desarrollo que dan lugar a la planta complementamente desarrollada, incluidas las plántulas y las semillas. Por su importancia, puede considerarse que un aspecto adicional de la invención es una semilla que comprende una molécula de ácido nucleico de la presente invención, es decir, una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica una delta-endotoxina, caracterizada por que la secuencia de nucleótidos que codifica una delta-endotoxina está representada por SEQ ID NO:1 o presenta al menos un 90% de identidad con la misma. Tanto en el caso de las plantas como en el de las semillas, se prefieren las plantas que sufren daños causados por áfidos (*Myzus persicae* en particular), o en las que se ha comprobado que dichos insectos facilitan la transmisión de virus. Así, entre otras, se tiene preferencia por plantas, partes de plantas o semillas de melocotonero (*Prunus persica*), albaricoquero (*Prunus armeniaca*), ciruelo (*Prunus domestica*), almendro (*Prunus dulcis*), otras especies del género *Prunus*, cítricos, patata, tomate y otras especies de la familia *Solanaceae*, tabaco, remolacha, remolacha azucarera, caña de azúcar, girasol, apio, mostaza, pimiento, calabaza, alcachofa,

espárrago, judías, brécol, coles de Bruselas, repollo, coliflor, col, col rizada, zanahoria, pepino, nabo, berenjena, lechuga, perejil, chirivía, guisante, alfalfa, pimiento, rábano, espinaca, berro, sandía, melón, maíz, sorgo, mijo, trigo, cebada, centeno, avena, arroz, soja, colza, cártamo, maní, batata, yuca, café, coco, piña, cacao, té, plátano, aguacate, higo, guayaba, mango, papaya, olivo, algodón, coníferas, azalea, hortensia, hibisco, rosas, tulipanes, narcisos, petunias, claveles, flor de pascua, y crisantemos.

Es también un aspecto de la presente invención cualquier polipéptido que comprenda la secuencia de la nueva delta-endotoxina de la presente invención o secuencias con un alto grado de identidad con la misma que puedan considerarse delta-endotoxinas. Así, la presente invención se refiere también a un polipéptido que comprende una secuencia de aminoácidos caracterizada por que dicha secuencia de aminoácidos:

- a) es la representada por SEQ ID NO:2;
- b) es la secuencia de una delta-endotoxina que está codificada por la secuencia representada por SEQ ID NO:1 o por una secuencia que presenta al menos un 90% de identidad con la misma, o
- c) es la secuencia de una delta-endotoxina que presenta al menos un 80% de identidad con la secuencia representada por SEQ ID NO:2.

Cualquiera de las secuencias que responde a la definición de a), b) o c) se considera una proteína (delta-endotoxina) de la presente invención. Se prefiere especialmente que la proteína presente tres dominios estructurales que se correspondan con los dominios N, M y C de las proteínas Cry. Muy especialmente, se prefiere que la secuencia de aminoácidos de la delta-endotoxina comprenda un dominio tipo ricina b.

En el caso de la alternativa c), el porcentaje de identidad puede ser, entre otros, siendo el 100% la opción a). Es una posible realización que el polipéptido se corresponda exactamente con el representado por SEQ ID NO:2.

En cualquiera de las opciones, se prefiere muy especialmente que el polipéptido tenga actividad insecticida contra al menos un hemíptero, preferiblemente al menos una especie de áfido, teniendo especial preferencia por que tenga actividad insecticida contra el áfido *Myzus persicae*.

En una posible realización, compatible con las anteriores, la secuencia de la delta-endotoxina de la presente invención está ligada a un segundo polipéptido, formando una proteína de fusión. Entre los posibles ejemplos de interés como segundo polipéptido

puede citarse la proteína de unión a la maltosa, como en el ejemplo 18 de la solicitud WO2010/099365. La delta-endotoxina de la presente invención también puede estar unida a otras secuencias de aminoácidos de posible interés, como puede ser una secuencia líder en 5' o, por ejemplo, una cola de histidinas que facilite su posterior
5 aislamiento.

Adicionalmente, es también un aspecto de la presente invención una composición que comprende un polipéptido de la presente invención según se ha definido en los párrafos previos, un vector de la presente invención o una célula hospedadora de la presente invención, composición que se considera una composición de la presente invención.
10 Preferiblemente, la composición comprende también un excipiente y/o un vehículo agronómicamente aceptable. La composición puede estar en diversas formas, preferiblemente formas adecuadas para que la composición pueda ser administrada como composiciones plaguicida, tales como polvo, granulado, emulsión, coloide, suspensión o solución. La suspensión o solución puede ser acuosa o puede ser que el
15 vehículo líquido sea, por ejemplo, un aceite o una emulsión.

Se prefiere que la composición comprenda un polipéptido de la presente invención o una célula hospedadora de la presente invención. En este segundo caso, como se ha comentado previamente, se tiene preferencia por las células hospedadoras bacterianas o de levadura, particularmente de células de una especie que se selecciona entre
20 *Pseudomonas spp.*, *Bacillus megaterium*, *Bacillus thuringiensis* Bt 4Q7 o *Saccharomyces cerevisiae*.

En la posible realización en la que la composición comprende un polipéptido de la presente invención, el mismo puede estar formando parte de una composición sólida, líquida o preparada para ser administrada en forma de aerosol. El polipéptido puede estar
25 en forma cristalina, de forma que formará una suspensión en el caso de que la composición se prepare en un vehículo acuoso, o puede estar solubilizado en un vehículo líquido, que puede ser acuoso.

Las composiciones de la presente invención, como se comentó previamente, pueden contener otros distintos componentes adicionales, entre ellos, otros diversos plaguicidas
30 o compuestos que incrementen su acción, incluidas las esporas de *Bacillus thuringiensis*, que pueden pertenecer a la cepa Bt 1.5H, a la cepa Bt 4Q7, a cualquier otra cepa natural, o a cepas recombinantes, incluidas cepas recombinantes que comprendan los vectores de la presente invención.

Puesto que uno de los principales objetivos de la presente invención es combatir las plagas de pulgón *Myzus persicae*, un aspecto adicional de la presente invención es el uso de las composiciones de la presente invención para combatir plagas de hemípteros, especialmente áfidos, y muy especialmente plagas del pulgón *Myzus persicae*.

- 5 Preferiblemente, el uso de la presente invención es para combatir plagas que produzcan daños en plantas, cultivos de plantas, o a los productos de la cosecha de los mismos. Se prefiere especialmente que las plantas a tratar sean plantas que sufren daños por infestaciones de *Myzus persicae* o plantas en las que dicho pulgón puede actuar como transmisor de virus. Como ya se comentó, estas plantas son principalmente el
- 10 melocotonero y otros frutales del género *Prunus*, así como cítricos, crucíferas, solánaceas e, incluso, coníferas y plantas ornamentales. Entre las especies a tratar se pueden citar melocotonero, albaricoquero, ciruelo, almendro, otras especies del género *Prunus*, cítricos, patata, tomate y otras especies de la familia *Solanaceae*, tabaco, remolacha, remolacha azucarera, caña de azúcar, girasol, apio, mostaza, pimiento,
- 15 calabaza, alcachofa, espárrago, judías, brécol, coles de Bruselas, repollo, coliflor, col, col rizada, zanahoria, pepino, nabo, berenjena, lechuga, perejil, chirivía, guisante, alfalfa, pimiento, rábano, espinaca, berro, sandía, melón, maíz, sorgo, mijo, trigo, cebada, centeno, avena, arroz, soja, colza, cártamo, maní, batata, yuca, café, coco, piña, cacao, té, plátano, aguacate, higo, guayaba, mango, papaya, olivo, algodón, coníferas, azalea,
- 20 hortensia, hibisco, rosas, tulipanes, narcisos, petunias, claveles, flor de pascua, y crisantemos. Se tiene especial preferencia por los frutales del género *Prunus*, por una parte, y por el tabaco y la patata, por la especial importancia que tiene en esta última para la transmisión de virus.

El uso comprende la administración de las composiciones de la invención por cualquier

25 método conocido, como ya se comentó previamente, apropiado según la formulación de la composición, tal como pulverización, administración en el agua de riego, depósito directo sobre órganos de la planta o junto a la base de la misma...

Otro aspecto más de la invención es un método para la producción de un polipéptido de la presente invención. El método comprende las etapas de:

- 30 a) obtener un organismo recombinante que exprese el polipéptido de la presente invención;
- b) cultivar dicho organismo recombinante;
- c) inducir, de ser necesario, la expresión del polipéptido en dicho organismo

recombinante;

d) purificar el polipéptido a partir del organismo recombinante.

El organismo recombinante puede ser, por ejemplo, una planta transgénica de la presente invención o una célula hospedadora de la presente invención u otro
5 microorganismo recombinante, preferiblemente una célula bacteriana o de levadura que o bien comprenda un vector de expresión de la presente invención, o bien haya integrado en su genoma una molécula de ácido nucleico de la presente invención, especialmente si en dicha molécula de ácido nucleico la secuencia de nucleótidos que codifica una delta-endotoxina estaba ya operativamente unida a un promotor que permite su expresión en
10 dicho microorganismo recombinante. Un ejemplo de purificación del polipéptido de la presente invención, en este caso la delta-endotoxina de SEQ ID NO:2, se presenta en el Ejemplo 2 de la presente invención, en el que se obtienen y cultivan células de *Escherichia coli* recombinante, transformadas con un plásmido que expresa dicha delta-endotoxina, induciendo la expresión de la delta-endotoxina con IPTG y procediendo a la
15 purificación a partir del cultivo, purificación que puede llevarse a cabo por cualquier medio conocido, como puede ser cromatografía de afinidad como en el Ejemplo 2 citado. En el caso de que la producción y purificación se produjera a partir de plantas transgénicas, las formas del polipéptido de la presente invención que comprenden proteínas de fusión o secuencias auxiliares de aminoácidos como colas de histidinas pueden ser de especial
20 utilidad para el aislamiento del polipéptido de la presente invención, pudiendo ser necesaria una etapa adicional en la que la secuencia auxiliar o el segundo polipéptido de la proteína de fusión se escinde, por ejemplo mediante el uso de una endoproteasa adecuada.

Finalmente, un aspecto adicional de la presente invención es un anticuerpo que se une
25 específicamente a un polipéptido de la presente invención. Preferiblemente, el anticuerpo se une específicamente al fragmento del polipéptido que constituye la delta-endotoxina de la presente invención, con especial preferencia por la realización en la que la secuencia de la delta-endotoxina es la representada por SEQ ID NO:2. En una realización preferida, el anticuerpo es un anticuerpo monoclonal.

30 La invención se explicará ahora con más detalle con ayuda de los Ejemplos y Figuras que aparecen a continuación.

Ejemplos

- Ejemplo 1: identificación y aislamiento del nuevo gen cry

1.1. Identificación y aislamiento del nuevo gen

El nuevo gen *cry* se identificó en la cepa de *Bacillus thuringiensis* Bt H1.5, una cepa perteneciente a la colección Bt de la Universidad Pública de Navarra, Laboratorio de Bioinsecticidas Microbianos (Iriarte et al., 1998; Iriarte et al., 2000: artículos que describen cómo se realizó la toma de muestras y se construyó la colección). La cepa Bt H1.5 en concreto fue aislada en la isla de El Hierro, Islas Canarias, España.

Esta cepa se cultivó en medio LB (Luria-Bertani) a 28 °C durante 16 horas a 200 rpm y su DNA total (genómico más plasmídico) se purificó con el kit Wizard® DNA purification kit (Promega), siguiendo las instrucciones del fabricante para el aislamiento de DNA desde bacterias Gram positivas, lo cual implica el tratamiento previo de las células bacterianas con lisozima a una concentración final de 10 mg/ml y una incubación de 30 minutos a 37°C.

El DNA total fue utilizado para generar una librería combinada ("pooled") según las instrucciones de Illumina, Inc. para la preparación de muestras con el sistema TruSeq™ (http://supportres.illumina.com/documents/documentation/chemistry_documentation/samplepreps_truseq/truseqsampleprep/truseq_sample_prep_pooling_guide_15042173_a.pdf), que fue secuenciado utilizando la tecnología HiSeq™ 2000 Sequencing System (Illumina® Sequencing) en modo de lectura simple, con un tamaño de lectura de 50 nucleótidos (GATC Biotech, Alemania).

Las secuencias obtenidas se analizaron utilizando Blast (Altschul et al., 1990) con bases de datos construidas con secuencias de toxinas Bt tales como Cry (cristalinas), Cyt (citolíticas), Vip1, Vip2 y Vip3 (proteínas insecticidas vegetativas), Sip (proteína insecticida secretables) ya descritas como así también aquellas producidas por la bacteria *Bacillus sphaericus* tales como Mtx 1, Mtx2 y Mtx3 (toxinas mosquitocidas) y binarias Bina/BinB (Berry, 2012), que fueron obtenidas de bases de datos públicas tales como Genbank (Benson et al., 2005). Alternativamente, también se utilizó el servidor de anotación automatizada RAST (Rapid Annotation using Subsystem Technology) (Aziz et al., 2008) el cual es un sistema totalmente automatizado para anotar genomas de bacterias y arqueas completos o casi completos, proporcionando anotaciones del genoma de alta calidad de todo el árbol filogenético (<http://rast.nmpdr.org>). Además de identificar las regiones codificantes o CDSs, identifica los genes de ARNr (ARN ribosomal) y ARNt (ARN de transferencia), asigna funciones a los genes y utiliza esta información para reconstruir la red metabólica. Este servidor permite además la descarga de los resultados en distintos formatos (GenBank, Fasta, GFF, etc) y la navegación en el genoma anotado

facilitando de esta manera el análisis comparativo con otros genomas anotados.

El nuevo gen *cry* fue identificado mediante la utilización de Blast (Blastx) y el ORF (Open Reading Frame: marco abierto de lectura) o secuencia codificante delimitado manualmente desde el codón de inicio (ATG) hasta el codón de stop (TAA) y con respecto a la secuencia homóloga de referencia con mayor cobertura y porcentaje de identidad. También se tuvo en cuenta la secuencia típica de unión a ribosoma o secuencia de Shine-Dalgarno, la cual está constituida por los siguientes nucleótidos 5'-AAGGAGG -3', previa al codon de inicio ATG y separada por 7 nucleótidos en la secuencia 5'-TTTATGC-3'. Esta secuencia de unión a ribosoma o secuencia de Shine-Dalgarno ya ha sido identificada en *Bacillus anthracis*, especie que está estrechamente relacionada a *Bacillus thuringiensis* (Steichen et al., 2003). La predicción de codones de inicio, ATG en nuestro caso, en base a la presencia de secuencias Shine-Dalgarno ya conocidas; minimiza significativamente la probabilidad de cometer error en la predicción de dicho codón (Starmer et al., 2006), especialmente, en genes como el de la presente invención, donde existen dos codones ATG próximos y separados por el nucleótido C (TTTATGCATG).

Este ORF, cuya secuencia está representada por SEQ ID NO:1, presentaba una longitud total de 2.397 bp (pares de bases), codifica una proteína de 799 aminoácidos, con un peso molecular aproximado de 89 kDa, secuencia de aminoácidos que está representada por SEQ ID NO:2.

1.2. Análisis bioinformático y comparativo de la secuencia de la proteína y de su secuencia codificante

La secuencia de aminoácidos de la nueva proteína presentó una identidad en Blast (BlastP) máxima del 76,56% y el 76,20% al compararlas contra otras proteínas previamente descritas en las bases de datos públicas, tanto comparando con secuencias "nr" (secuencias de proteínas no redundantes), como comparando con secuencias "pat" (secuencias de proteínas patentadas), respectivamente. Puesto que las proteínas con las que presentaba mayor porcentaje de identidad eran todas ellas proteínas Cry o delta-endotoxinas, la proteína de la presente invención es una proteína Cry nueva.

Las Tablas 2 y 3 muestran los resultados obtenidos al realizar el BlastP mediante los recursos de uso público accesibles para tal fin en la página web del National Center for Biotechnology Information (NCBI) de EE.UU. (http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LI

NK LOC=blasthome), tanto seleccionando la base de datos de proteínas no redundantes (nr) (Tabla2), como seleccionando la base de datos de proteínas patentadas (pat) (Tabla 3). La Tabla 4, por su parte, muestra los resultados obtenidos con las herramientas análogas disponibles en la página web del EBI (European Bioinformatics Institute: Instituto Europeo de Bioinformática) para la consulta del ENA (European Nucleotide Archive: Archivo Europeo de Nucleótidos) (<http://www.ebi.ac.uk/ena/>), realizando la consulta sobre la secuencia de nucleótidos de SEQ ID NO:1.

Tabla 2
Lista de proteínas no redundantes con mayor porcentaje de identidad con respecto a la delta-endotoxina de la presente invención

Secuencia localizada S	Núm. Acceso en Genbank	% Identidad	% Positivos	Longitud de alineamiento	Posiciones discrepantes	Espacios abiertos	Punto de inicio Q ^a	Punto final Q ^b	Punto inicio S ^c	Punto final S ^d	Valor e	Valor de bits
Proteína pesticida	AGU13869	76,56	83,24	704	154	5	1	700	24	720	0.0	1055
Proteína pesticida	AGU13855	76,20	83,14	706	156	6	1	700	24	723	0.0	1049
Proteína similar a Cry41Aa1	AEH76822.1	57,91	70,92	815	313	15	1	799	1	801	0.0	879
Hipotética Proteína	EJR94916.1	56,73	68,45	802	312	13	12	799	12	792	0.0	841
Hipotética Proteína	EJR93828.1	42,08	55,20	846	420	22	1	799	1	823	6,00E-179	546
Hipotética Proteína	EJR29963.1	40,68	55,81	826	447	20	1	799	1	810	4,00E-165	510
Proteína Cry que mata células cancerosas	BAD35160.1	39,19	54,30	860	423	23	1	796	1	824	2,00E-152	478
Proteína Cry	BAD35166.1	39,64	53,94	825	457	21	1	799	1	810	2,00E-148	467
Proteína Cry que mata células cancerosas	EEM68667.1	39,24	53,36	864	435	24	1	799	26	864	4,00E-146	462
Proteína cristalina CryE6S	BAB78603.1	37,94	54,74	738	401	20	1	713	1	706	2,00E-121	406
Proteína Cry que mata células cancerosas	BAD35163.1	48,56	65,35	381	176	8	1	365	1	377	2,00E-89	310
Cry32	BAD35163.1	39,60	54,34	346	180	12	471	796	492	828	1,00E-49	196

^a Punto de inicio de la comparación en la secuencia de búsqueda (delta-endotoxina de la invención)

^b Punto final de la comparación en la secuencia de búsqueda (delta-endotoxina de la invención)

^c Punto de inicio de la comparación en la secuencia localizada

^d Punto final de la comparación en la secuencia localizada

Tabla 3

Lista de proteínas de documentos de patente con mayor porcentaje de identidad con respecto a la delta-endotoxina de la presente invención

Secuencia localizada S	Núm. Acceso en Genbank	% Identidad	% Positivos	Longitud alineamiento	Posiciones discrepantes	Espacios abiertos	Punto inicio Q ^a	Punto final Q ^b	Punto inicio S ^c	Punto final S ^d	Valor e	Valor de bits
Secuencia 87 de US8318990	AGA40054.1	76,20	83,14	706	156	6	1	700	24	723	0,0	1049
Secuencia 84 de US8461421	AGP18038.1	57,91	70,92	815	313	15	1	799	6	806	0,0	879
Secuencia 94 de US8461421	AGP18048.1	38,15	51,81	886	429	25	1	796	1	857	3,00E-154	482
Secuencia 140 de US8461421	AGP18084.1	39,24	53,36	864	435	24	1	799	26	864	9,00E-148	462
Secuencia 70 de US8461421	AGP18024.1	39,82	55,72	673	334	20	12	650	12	647	6,00E-118	392
Secuencia 105 de US8461421	AGP18059.1	36,57	52,71	700	374	20	1	650	3	682	9,00E-118	378
Secuencia 92 de US8461421	AGP18046.1	33,38	47,04	710	365	20	42	656	50	746	3,00E-86	304
Secuencia 68 de US8461421	AGP18022.1	33,53	49,06	695	387	22	1	650	14	678	9,00E-84	296
Secuencia 82 de US8318990	AGA40049.1	33,88	47,11	726	346	28	12	656	12	684	9,00E-84	296
Secuencia 91 de US8461421	AGP18045.1	33,56	46,77	727	347	29	12	656	12	684	3,00E-80	285

^a Punto de inicio de la comparación en la secuencia de búsqueda (delta-endotoxina de la invención)

^b Punto final de la comparación en la secuencia de búsqueda (delta-endotoxina de la invención)

^c Punto de inicio de la comparación en la secuencia localizada

^d Punto final de la comparación en la secuencia localizada

Tabla 4

Lista de secuencias de nucleótidos con mayor porcentaje de identidad con respecto a la secuencia codificante de la delta-endotoxina de la presente invención

Secuencia localizada S	Núm. Acceso en EMBL-Bank	% identidad	Longitud de alineamiento	Punto de inicio Q ^a	Punto final Q ^b	Punto inicio S ^c	Punto final S ^d	Valor e	Valor de bits
Gen de proteína plaguicida de <i>Bacillus thuringiensis</i> , cepa ARP242: cds	KC156704	87	1758	1	1758	70	1827	0	2063
Secuencia 38 de US8318900 / WO 2010099365 / JP 2012519000-A	GZ600762 / HH806253 / HV951516	87	1757	1	1757	70	1829	0	2056
Gen de proteína plaguicida de <i>Bacillus thuringiensis</i> , cepa ARP256: cds	KC156680	87	1757	1	1757	1	1829	0	2056
Contig1.62 de secuenciación del genoma completo de <i>Bacillus cereus</i> VDM022	AHFP01000062	73	1345	1	1345	6436	5086	9E-136	500
Supercontig1.5 de asociación de fragmentos de secuenciación del genoma completo de <i>Bacillus cereus</i> VDM022	JH791841	73	1345	1	1345	67926	66576	9E-136	500
Secuencia 170 de WO2010099365 / Secuencia 123 de JP 2012519000-A	HH806385 / HV951601	72	1585	1	1585	1	1585	1E-130	483
Secuencia 24 de WO2009158470 / Secuencia 24 de US8461421 / Secuencia 24 de JP2011526150-A	HC309841 / HJ226143 / HV549524	71	1345	1	1345	16	1360	2E-101	385
Gen de proteína similar a Cry41Aa1 de <i>Bacillus thuringiensis</i> , cepa Sb1029: cds	JF521582	71	1345	1	1345	16	1360	2E-101	385
Secuencia 169 de WO2010099365 / Secuencia 122 de JP 2012519000-A	HH806384 / HV951600	70	1585	1	1585	1	1585	1E-87	483
Secuencia 24 de WO2009158470 / Secuencia 24 de US8461421 / Secuencia 24 de JP2011526150-A	HC309841 / HJ226143 / HV549524	76	215	1372	1586	1375	1592	4E-21	119
Gen de proteína similar a Cry41Aa1 de <i>Bacillus thuringiensis</i> , cepa Sb1029: cds	JF521582	76	215	1372	1586	1375	1592	4E-21	119

Secuencia localizada S	Núm. Acceso en EMBL-Bank	% Identidad	Longitud de alineamiento	Punto de inicio Q ^a	Punto final Q ^b	Punto inicio S ^c	Punto final S ^d	Valor e	Valor de bits
Contig1.62 de secuenciación del genoma completo de <i>Bacillus cereus</i> VDM022	AHFP01000062	75	221	1372	1592	5071	4848	6E-16	102
Supercontig1.5 de asociación de fragmentos de secuenciación del genoma completo de <i>Bacillus cereus</i> VDM022	JH791841	75	221	1372	1592	33561	66338	6E-16	102
Gen de proteína plaguicida de <i>Bacillus thuringiensis</i> , cepa ARP277: cds	KC156705	70	455	1497	1951	1545	2005	2E-70	73
Gen de hipotética proteína: proteína Cry parasporina-3 que mata células cancerosas, mitad C-terminal, cds completa, clon: GC-I	AB116649	74	152	1874	2025	4903	5057	6E-4	62
Gen de hipotética proteína: proteína Cry parasporina-3 que mata células cancerosas, mitad C-terminal, cds completa, clon: GC-II	AB116650	74	152	1874	2025	5367	5521	6E-4	62

^a Punto de inicio de la comparación en la secuencia de búsqueda (gen de la delta-endotoxina de la invención)

^b Punto final de la comparación en la secuencia de búsqueda (gen de la delta-endotoxina de la invención)

^c Punto de inicio de la comparación en la secuencia localizada

^d Punto final de la comparación en la secuencia localizada

E software Geneious Pro v6.1.7 (Drummond et al., 2014) y su herramienta Geneious Alignment se utilizó para realizar el alineamiento múltiple de la secuencia de la proteína CryBM1 contra otras proteínas Cry conocidas, para las cuales se ha descrito que presentan actividad contra áfidos, incluida la proteína representada por la secuencia con número de identificación 205 en la patente US8318900B2 y en la solicitud internacional WO2010099365 (a la que se aludirá de forma abreviada en lo sucesivo como Seq205-US8318900). Los resultados de dicho alineamiento se muestran en la Fig. 1. Este alineamiento se utilizó posteriormente para generar el dendrograma representativo de la relación filogenética entre las diversas proteínas Cry activas contra áfidos que se representa en la Fig. 2 con la herramienta Geneious Tree Builder y obtener una matriz de % de identidad de cada una de ellas (Tabla 5).

Tabla 5
Porcentaje de similitud entre proteínas Cry activas contra áfidos,
incluida la proteína Cry de la invención (CryBM1)

	Cry2Aa	Cry3Aa	Cry4Aa	Cry11Aa	CryBM1	Seq205-US8318900
Cry2Aa	100	15,5	13,4	15,7	14,2	13,9
Cry3Aa	15,5	100	25,2	11,6	26,3	17,1
Cry4Aa	13,4	25,2	100	9,1	16,5	11,8
Cry11Aa	15,7	11,6	9,1	100	11,6	12,7
CryBM1	14,2	26,3	16,5	11,6	100	15,0
Seq205-US8318900	13,9	17,1	11,8	12,7	15,0	100

15

Como puede verse en la Fig. 1, la nueva proteína Cry, al igual que las demás (excepto Cry2Aa, Cry11Aa y Seq205-US8318900), resultó ser una proteína Cry con tres dominios N, M y C, los dominios que se consideran responsable de la actividad insecticida contra distintas especies de insectos.

20 La nueva proteína Cry de la presente invención (CryBM1), posee, sin embargo, una característica singular representada por un fragmento de secuencia tipo dominio conservado de ricina (numero acceso InterPro IPR000772) constituido por 148

aminoácidos localizado en el extremo carboxilo terminal entre las posiciones 652 a 799 y representada por la secuencia TPTPEPVVDGIYQIVTALNNSVAENGGPTTRGGPDQVKLSPFYNSTDQKWEFIYDSNE DVYTIRNLAGGFLTYFMLNPGYPVLAIRPQWATENQKWIVEPAGNGYYYLRKSKSVPTAAA FAPNAADGSVVRMSAVDFSTNOKFKLNKL; más concretamente, un dominio del tipo de la cadena B de la ricina natural, que es la cadena que tiene actividad como lectina (es decir, una proteína que se une a hidratos de carbono), pues se une a residuos terminales de galactosa presente en las superficies de las células. La ricina natural que se produce en las semillas de las plantas de ricino (*Ricinus communis*) es una proteína heterodimérica, compuesta por una cadena A y una cadena B, que resulta ser una lectina altamente tóxica para una amplia variedad de organismos (Chougule and Bonning, 2012). Sin querer limitarse a ninguna teoría específica, se propone que ese dominio de ricina podría ser el responsable de la elevada toxicidad de la proteína Cry de la presente invención contra *M. persicae* que se demuestra en ejemplos posteriores de la presente invención.

- Ejemplo 2: Clonación del gen y producción de la proteína recombinante

Una vez identificado el nuevo gen *cry*, fue amplificado mediante PCR (*Polymerase Chain Reaction*) utilizando una *Taq* polimerasa a prueba de errores (PrimeSTAR HS DNA polymerase, Takara) con un cebador directo que permitía añadir una diana de restricción *NcoI* (5'-CCATGGATGAACCAAAATTATAACAACAATG-3') (SEQ ID NO:5) inmediatamente antes del codón de inicio (ATG) y un cebador reverso que permitía añadir una diana de restricción *Safl* (5'-GTCGACTAACTTATTCAGTTTGAATTTTTGATTTG -3') (SEQ ID NO:6), detallados en cursiva en la secuencia. El cebador reverso además carecía el codón de parada (TAA) permitiendo la fusión de la proteína recombinante a la cola de 6 histidinas codificada en el vector. Las reacciones se llevaron a cabo en un volumen final de 25 µl con 5 µl de *buffer* de reacción 5X, 10 mM de dNTPs, 6,3 pmol de cada cebador, 0,5 U de la DNA polimerasa a prueba de errores PrimeSTAR HS DNA polymerase (Takara) y 100 ng de DNA total. Cada reacción de llevó a cabo en las siguientes condiciones: 4 min. de desnaturalización inicial a 94°C, 35 ciclos de amplificación (desnaturalización de 1 min. desnaturalización a 94°C, 1 min. de *annealing* (apareamiento) a 52°C y 2,5 min. de extensión a 72°C) y con un paso de extensión final a 72°C durante 10 min.

El fragmento amplificado fue ligado al vector pJET (Thermo Scientific) siguiendo las instrucciones provistas por el fabricante, separado del vector de clonado mediante

digestión con las enzimas de restricción *Nco*I y *Sal*I y purificado del gel de agarosa con el kit Nucleospin Gel and PCR Cleanup (Macherey-Nagel). El fragmento purificado se subclonó en el vector de expresión pET-28b(+) (Novagen) (ver Fig. 3) predigerido con las enzimas de restricción *Nco*I y *Sal*I y siguiendo técnicas rutinarias de biología molecular descritas por Sambrook y Russell (Sambrook and Russell, 2001). Este vector, posee un promotor fuerte para la RNA polimerasa viral del Fago T7 y un gen de resistencia a al antibiótico Kanamicina para facilitar su selección

Una vez obtenido el vector recombinante de expresión, representado en el Fig. 3, el mismo se utilizó para transformar células de *Escherichia coli* BL21(DE3) (Life Technologies), siguiendo técnicas rutinarias de biología molecular descritas por Sambrook y Russell (Sambrook and Russell, 2001). Estas células portan en su genoma el gen codificante de la RNA polimerasa viral del fago T7 inducible por IPTG (isopropil- β -D-1-tiogalactopiranosido), la cual reconoce el promotor codificado en el vector de expresión pET-28b(+) (Novagen) y transcribe el gen clonado llevando a la producción de la proteína codificada por tal gen.

La producción de la proteína recombinante CryBM1 se llevó a cabo en las células de *Escherichia coli* obtenidas, induciendo su expresión con 1mM (concentración final) de IPTG a 37 °C y una 200 rpm de agitación durante 16 hs.

La purificación de la proteína recombinante portadora de una cola de 6 histidinas se llevó a cabo mediante cromatografía de afinidad en columnas de níquel con TED (tris-carboximetil-etilen-diamina) como ligando quelante, con el kit Protino Ni-Ted 1000 (Macherey-Nagel), siguiendo las instrucciones del fabricante. La concentración de la proteína soluble purificada fue determinada por el método de Bradford (Bradford, 1976). La nueva proteína Cry producida en *E. coli* se obtuvo en fase soluble a una concentración aproximada de 1 mg/ml a partir de un litro de medio de cultivo. Al someter a electroforesis en gel SDS-PAGE una alícuota de la muestra purificada obtenida, se observó una banda a una altura correspondiente a un tamaño de 89 kDa, equiparable al peso molecular teórico de la secuencia de la proteína recombinante y que se muestra en la Fig. 4.

- Ejemplo 3: Análisis de la actividad insecticida

El análisis de la actividad insecticida se llevó a cabo mediante una simplificación de la metodología de bioensayos descrita por Sadeghi et al. (Sadeghi et al., 2009) utilizando ninfas de *M. persicae* de primer estadio (N1) suministradas por el insectario del centro de investigación de los presentes inventores. Brevemente, la proteína soluble fue diluida en dieta artificial para áfidos la cual estaba constituida por sacarosa al 20% P/V en agua

destilada estéril. Cada dilución se incorporó entre dos láminas de Parafilm® "M" (Bemis Flexible Packaging, Nena, Wiscosin, EE.UU.) y cubriendo cajas circulares pequeñas de 3 cm de diámetro (1,5 cm de alto) sin tapa. Cada caja contenía 15 ninfas N1. De esta manera, las ninfas eran capaces de alimentarse a través de la primer membrana de Parafilm® "M" y directamente sobre la mezcla de dieta artificial+toxina en los tratamientos o simplemente de la dieta artificial en el caso del control negativo. La Fig. 5 muestra fotografías de las cajas empleadas, cubiertas con Parafilm® "M" y con la gota de dieta artificial entre dos capas de dicho material.

Los bioensayos se llevaron a cabo en cámara bajo las siguientes condiciones controladas: 25°C de temperatura, 70% de humedad relativa y bajo fotoperíodo de 16:8 horas de luz/oscuridad, respectivamente. La mortalidad producida se registró a las 72 horas.

En un principio, se realizaron ensayos preliminares en dos repeticiones y dos concentraciones de 100 y 1000 µg/ml de la proteína Cry de la presente invención, CryBM1. Estos ensayos permitieron determinar las concentraciones aproximadas que producirían una mortalidad de entre el 5 al 95%. La concentración letal media (CL₅₀) se determinó entonces a cuatro concentraciones de 20, 25, 32 y 40 µg/ml y a tres repeticiones por cada concentración y el control negativo.

Los resultados obtenidos se muestran a continuación en la Tabla 6:

20

Tabla 6

Porcentajes de mortalidad producidos por distintas concentraciones de CryBM1 al ser ingeridas por ninfas neonatas de *Myzus persicae*

Concentración (µg/ml)	Insectos tratados	Insectos muertos	% de mortalidad
0	120	15	12,5
20	126	20	15,9
25	125	29	23,2
32	119	65	54,6
40	120	100	83,3

La CL₅₀ fue estimada a partir de los datos anteriores, utilizando el programa estadístico Polo PC (LeOra-Software, 1987) basado en el análisis probit de Finney (1971). El valor

25

obtenido y los detalles del análisis probit en los que se basa se muestran a continuación en la Tabla 7.

Tabla 7

5 Concentración letal media (CL_{50}) estimada para la nueva proteína CryBM1 contra ninfas neonatas de *Myzus persicae*

Tratamiento	LC_{50} ($\mu\text{g/ml}$)	Regresión lineal		Bondad de ajuste	
		Pendiente \pm EE*	$a^*\pm$ EE*	χ^2	g.l.*
Proteína Cry	32,7	10,3 \pm 1,4	-10,6 \pm 2,1	0,55	2

*a: intersección con la recta de regresión

EE: error estándar

g.l.: grados de libertad

- 10 La CL_{50} determinada fue de tan solo 32,7 $\mu\text{g/ml}$. Este valor puede considerarse la CL_{50} más baja descrita hasta la fecha para una proteína Cry de Bt aplicada a una especie de hemíptero plaga.

15

BIBLIOGRAFÍA

- Ayres, N.M., Park W.D, 1994. Genetic transformation of rice. *Critical Reviews in Plant Science* 13:219-239.
- Altschul, S.F., et al., 1990. Basic local alignment search tool. *J Mol Biol* 215, 403-410.
- 5 Altschul, S.F., Madden T.L., Schäffer A.A., Zhang J., Zhang Z., Miller W., Lipman D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acid Res.* 25:3389.
- Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R.A., McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A., Zagnitko, O., 10 2008. The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9, 75.
- Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Wheeler, D.L., 2005. GenBank. 15 *Nucleic Acids Research* 33, D34-D38.
- Bommineni, V.M., Jauhar P.P, 1997, "an evaluation of target-cells and tissues used in genetic-transformation of cereals", *Maydica*, 42(2), 107-120.
- Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* 72, 248-254.
- 20 Chougule, N.P., Bonning, B.C., 2012. Toxins for transgenic resistance to hemipteran pests. *Toxins (Basel)* 4, 405-429.
- Diaz, B.M., Biurrun, R., Moreno, A., Nebreda, M., Ferreres, A., 2006. Impact of ultraviolet-blocking plastic films on insect vectors of virus diseases infesting crisp lettuce. 25 *Hortscience* 41, 711-716.
- Drummond, A. J., et al., 2014. Geneious Pro v6.1.7. [Online.]. <http://www.geneious.com/>. 2014.
- Dulmage, H.T., 1970. Insecticidal Activity of Hd-1, a New Isolate of *Bacillus thuringiensis* Var. Alesti. *Journal of Invertebrate Pathology* 15, 232-&.
- 30 Estévez J, Leiro J, Santamarina MT, Domínguez J, Ubeira FM. Monoclonal antibodies to turbot (*Scophthalmus maximus*) immunoglobulins: characterization and applicability in immunoassays. *Vet Immunol Immunopathol* 1994; 41: 353-366.
- Estruch, J.J., Warren, G.W., Mullins, M.A., Nye, G.J., Craig, J.A., Koziel, M.G., 1996. Vip3A, a novel *Bacillus thuringiensis* vegetative insecticidal protein with a wide 35 spectrum of activities against lepidopteran insects. *Proceedings of the National Academy of Sciences of the United States of America* 93, 5389-5394.
- Estruch, J.J., Yu, C.G., 1998. Plant pest control. Patent WO 9844137.
- Federici, B.A., Park, H.W., Bideshi, D.K., Wirth, M.C., Johnson, J.J., 2003. Recombinant bacteria for mosquito control. *J Exp Biol* 206, 3877-3885.
- 40 Goldberg, L.J.a.J.M., 1977. A bacterial spore demonstrating rapid larvicidal activity against *Anopheles sergentii*, *Uranotaenia unguiculata*, *Culex univittatus*, *Aedes aegypti* and *Culex pipiens*. *Mosquito News*. 37.

- Harrington, H.F.v.E.a.R., 2007. Aphids as crop pests. CAB International 2007.
- Hellens R, Mullineaux P., Klee H, 2000. Technical Focus: a guide to Agrobacterium binaty Ti vectors. Trends in Plant Science 5:446-451.
- Iglesias R, Leiro J, Santamarina MT, Sanmartín ML, Ubeira FM. Monoclonal antibodies
5 against diagnostic Anisakis simplex antigens. Parasitol Res. 1997; 83:755-761.
- Iriarte, J., Bel, Y., Ferrandis, M.D., Andrew, R., Murillo, J., Ferre, J., Caballero, P., 1998. Environmental distribution and diversity of *Bacillus thuringiensis* in Spain. Systematic and Applied Microbiology 21, 97-106.
- Iriarte, J., Porcar, M., Lecadet, M.M., Caballero, P., 2000. Isolation and characterization of
10 *Bacillus thuringiensis* strains from aquatic environments in Spain. Current microbiology 40, 402-408.
- Krieg, A., Huger, A.M., Langenbruch, G.A., Schnetter, W., 1983. *Bacillus thuringiensis* var. *Tenebrionis*, a New Pathotype Effective against Larvae of Coleoptera. Zeitschrift Fur Angewandte Entomologie-Journal of Applied Entomology 96, 500-508.
- 15 LeOra-Software, 1987. POLO-PC: a user's guide to Probit or Logit analysis. LeOra Software, Berkeley, Calif.
- Nebreda, M., Moreno, A., Perez, N., Palacios, I., Seco-Fernandez, V., Fereres, A., 2004. Activity of aphids associated with lettuce and broccoli in Spain and their efficiency as vectors of Lettuce mosaic virus. Virus Res 100, 83-88.
- 20 Ohba, M., Mizuki, E., Uemori, A., 2009. Parasporin, a new anticancer protein group from *Bacillus thuringiensis*. Anticancer Research 29, 427-433.
- Porcar, M., Grenier, A.M., Federici, B., Rahbe, Y., 2009. Effects of *Bacillus thuringiensis* delta-endotoxins on the pea aphid (*Acyrtosiphon pisum*). Applied and Environmental Microbiology 75, 4897-4900.
- 25 Sadeghi, A., Van Damme, E.J., Smagghe, G., 2009. Evaluation of the susceptibility of the pea aphid, *Acyrtosiphon pisum*, to a selection of novel biorational insecticides using an artificial diet. Journal of Insect Science (Ludhiana) 9, 1-8.
- Sambrook, J., Russell, D., 2001. Molecular cloning: a laboratory manual. 3rd ed. Cold Spring Harbor laboratory Press, Cold Spring Harbor, N.Y.
- 30 Sattar, S., Maiti, M.K., 2011. Molecular characterization of a novel vegetative insecticidal protein from *Bacillus thuringiensis* effective against sap-sucking insect pest. Journal of Microbiology Biotechnology 21, 937-946.
- Slater, R., Paul, V.L., Andrews, M., Garbay, M., Camblin, P., 2012. Identifying the presence of neonicotinoidresistant peach-potato aphid (*Myzus persicae*) in the
35 peach-growing regions of southern France and northern Spain. Pest Manag Sci 68, 634-638.
- Sulzer, J.H., Linne, C.v., Römer, J.J., Schellenberg, J.R. 1776. Abgekürzte Geshichte der Insecten nach dem Linaeischen System. Winterthur, bey H. Steiner u. Comp. Buchh., Schweiz.

REIVINDICACIONES

1. Una molécula de ácido nucleico aislada que comprende una secuencia de nucleótidos que codifica una delta-endotoxina, caracterizada por que la secuencia de nucleótidos que codifica una delta-endotoxina está representada por SEQ ID NO:1 o presenta al menos un 90% de identidad con la misma.

2. Molécula de ácido nucleico según la reivindicación 1, en la que la secuencia de nucleótidos que codifica una delta-endotoxina presenta al menos un porcentaje de identidad con la secuencia representada por SEQ ID NO:1 que se selecciona del grupo de 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5%, 99,6%, 99,7%, 99,8%, 99,9%.

3. Molécula de ácido nucleico según la reivindicación 1 ó 2, en la que la secuencia de nucleótidos que codifica una delta-endotoxina que comprende un fragmento de secuencia que codifica un dominio tipo ricina b que presenta al menos un 90% de identidad con la secuencia codificante de los aminoácidos 652-799 de SEQ ID NO:2.

4. Molécula de ácido nucleico según una cualquiera de las reivindicaciones anteriores, que comprende una secuencia de nucleótidos que codifica una delta-endotoxina que tiene actividad contra al menos un insecto hemíptero.

5. Molécula de ácido nucleico según la reivindicación 4, que comprende una secuencia de nucleótidos que codifica una delta-endotoxina que tiene actividad contra al menos una especie de áfido.

6. Molécula de ácido nucleico según la reivindicación 5, que comprende una secuencia de nucleótidos que codifica una delta-endotoxina que tiene actividad contra el áfido *Myzus persicae*.

7. Molécula de ácido nucleico según la reivindicación 1, cuya secuencia de nucleótidos está representada por SEQ ID NO:1.

8. Molécula de ácido nucleico según una cualquiera de las reivindicaciones 1 a 6, en la que la secuencia de nucleótidos que codifica una delta-endotoxina está operativamente unida a un promotor.

9. Molécula de ácido nucleico según la reivindicación 8, en la que el promotor se selecciona del grupo de un promotor capaz de dirigir la expresión de dicha secuencia de nucleótidos en una célula de planta, en una bacteria que se selecciona del grupo de
5 *Escherichia coli*, *Bacillus thuringiensis*, *Pseudomonas spp.*, o en una célula de *Saccharomyces cerevisiae*.

10. Molécula de ácido nucleico según la reivindicación 8 ó 9, que adicionalmente está operativamente unido a una secuencia de control adicional al promotor.

10

11. Un vector en el que está inserta una molécula de ácido nucleico de una cualquiera de las reivindicaciones 1 a 10.

12. Vector según la reivindicación 11, que es un plásmido.

15

13. Vector según la reivindicación 12, que es un vector de expresión que permite la expresión en una bacteria de la secuencia de nucleótidos que codifica una delta-endotoxina representada por SEQ ID NO:1 o de una secuencia codificante que presenta al menos un 90% de identidad con la misma.

20

14. Vector según la reivindicación 12, en el que la bacteria se selecciona del grupo de *Escherichia coli*, *Pseudomonas spp.*, *Bacillus megaterium* o *Bacillus thuringiensis*.

15. Vector según la reivindicación 11 ó 12, que es un vector de transformación de
25 plantas.

16. Vector según la reivindicación 15, en el que la secuencia de nucleótidos que codifica la delta-endotoxina representada por SEQ ID NO:1 o la secuencia codificante que presenta al menos un 90% de identidad con la misma está operativamente unida a
30 un promotor capaz de dirigir la expresión de dicha secuencia de nucleótidos en una célula de planta.

17. Vector según una cualquiera de las reivindicaciones anteriores, en que la secuencia codificante de la delta-endotoxina representada por SEQ ID NO:1 o la
35 secuencia codificante que presenta al menos un 90% de identidad con la misma está unida en fase de lectura a la secuencia codificante de un segundo polipéptido.

18. Una célula hospedadora que comprende un vector según una cualquiera de las reivindicaciones 11 a 17.

5 19. Célula hospedadora según la reivindicación 18, que es una célula de levadura.

20. Célula hospedadora según la reivindicación 18, que es una célula de bacteria.

10 21. Célula hospedadora según la reivindicación 20, que es una célula de una bacteria que se elige del grupo de *Escherichia coli*, *Pseudomonas spp.*, *Bacillus megaterium* o *Bacillus thuringiensis*.

22. Célula hospedadora según la reivindicación 18, que es una célula de planta.

15 23. Una planta transgénica que comprende la célula hospedadora de la reivindicación 22.

24. Una semilla transgénica que comprende una molécula de ácido nucleico de una cualquiera de las reivindicaciones 1 a 9.

20

25. Planta transgénica según la reivindicación 23 o semilla transgénica según la reivindicación 24, que se selecciona del grupo de melocotonero, albaricoquero, ciruelo, almendro, otras especies del género *Prunus*, cítricos, patata, tomate y otras especies de la familia *Solanaceae*, tabaco, remolacha, remolacha azucarera, caña de azúcar, girasol, apio, mostaza, pimiento, calabaza, alcachofa, espárrago, judías, brécol, coles de Bruselas, repollo, coliflor, col, col rizada, zanahoria, pepino, nabo, berenjena, lechuga, perejil, chirivía, guisante, alfalfa, pimiento, rábano, espinaca, berro, sandía, melón, maíz, sorgo, mijo, trigo, cebada, centeno, avena, arroz, soja, colza, cártamo, maní, batata, yuca, café, coco, piña, cacao, té, plátano, aguacate, higo, guayaba, mango, papaya, olivo, algodón, coníferas, azalea, hortensia, hibisco, rosas, tulipanes, narcisos, petunias, claveles, flor de pascua, y crisantemos.

30

26. Un polipéptido que comprende una secuencia de aminoácidos caracterizada por que dicha secuencia de aminoácidos:

35

a) es la representada por SEQ ID NO:2;

b) es la secuencia de una delta-endotoxina que está codificada por la secuencia representada por SEQ ID NO:1 o por una secuencia que presenta al menos un 90% de identidad con la misma, o

5 c) es la secuencia de una delta-endotoxina que presenta al menos un 80% de identidad con la secuencia representada por SEQ ID NO:2.

27. Polipéptido según la reivindicación 26, en el que la secuencia de aminoácidos representada por SEQ ID NO:2 o la delta-endotoxina definida en los apartados b) ó c) de la reivindicación 26 presenta tres dominios estructurales que se corresponden con los
10 dominios N, M y C de las proteínas Cry.

28. Polipéptido según la reivindicación 26 ó 27, que comprende una secuencia de aminoácidos que es la secuencia de una delta-endotoxina que presenta al menos un 80% de identidad con la secuencia representada por SEQ ID NO:2, en el que el
15 porcentaje de identidad se selecciona del grupo de 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5%, 99,6%, 99,7%, 99,8%, 99,9%.

29. Polipéptido según la reivindicación 26, cuya secuencia está representada por SEQ ID NO:2.

20

30. Polipéptido según una cualquiera de las reivindicaciones 26 a 29, que presenta actividad insecticida contra al menos un hemíptero.

31. Polipéptido según la reivindicación 30, que presenta actividad insecticida
25 contra al menos un áfido.

32. Polipéptido según la reivindicación 31, que presenta actividad insecticida contra *Myzus persicae*.

30 33. Polipéptido según una cualquiera de las reivindicaciones 26 a 32, en el que la secuencia de aminoácidos representada por SEQ ID NO:2 o que corresponde a una delta-endotoxina codificada por una secuencia que presenta al menos un 90% de identidad con SEQ ID NO:1 o que corresponde a una delta-endotoxina que presenta al menos un 80% de identidad con la secuencia representada por SEQ ID NO:2, está ligada
35 a un segundo polipéptido, formando una proteína de fusión.

34. Una composición que comprende un polipéptido de una cualquiera de las reivindicaciones 26 a 33, un vector de una cualquiera de las reivindicaciones 11 a 17 y/o una célula hospedadora de una cualquiera de las reivindicaciones 18 a 22.

5

35. Composición según la reivindicación 34, que adicionalmente comprende un excipiente agronómicamente aceptable y/o un vehículo agronómicamente aceptable.

36. Composición según la reivindicación 34 ó 35, que comprende una célula hospedadora de una cualquiera de las reivindicaciones 19 a 21.

10

37. Composición según la reivindicación 36, en la que las células hospedadoras se seleccionan entre las células de las especies *Pseudomonas spp.*, *Bacillus megaterium*, *Bacillus thuringiensis* Bt 4Q7 o *Saccharomyces cerevisiae*.

15

38. Composición según la reivindicación 34 ó 35, que comprende un polipéptido de una cualquiera de las reivindicaciones 26 a 33, en forma cristalina o solubilizada en un vehículo líquido.

20

39. Composición según una cualquiera de las reivindicaciones 34 a 38, que está en forma de polvo, granulado, emulsión, coloide, suspensión o solución.

40. Composición según una cualquiera de las reivindicaciones 34 a 39, que comprende un compuesto plaguicida adicional, un compuesto que incremente la acción plaguicida o esporas de *Bacillus thuringiensis*.

25

41. Uso de una composición de una cualquiera de las reivindicaciones 34 a 40 para combatir plagas de al menos un hemíptero.

30

42. Uso según la reivindicación 41, en el que el hemíptero es un áfido.

43. Uso según la reivindicación 42, en el que el áfido es *Myzus persicae*.

35

44. Uso según una cualquiera de las reivindicaciones 41 a 43, en el que la plaga a combatir afecta a una planta, un cultivo de plantas, o a los productos de la cosecha de los mismos.

45. Uso según la reivindicación 44, en el que la planta se selecciona del grupo de melocotonero, albaricoquero, ciruelo, almendro, otras especies del género *Prunus*, cítricos, patata, tomate y otras especies de la familia *Solanaceae*, tabaco, remolacha, 5 remolacha azucarera, caña de azúcar, girasol, apio, mostaza, pimienta, calabaza, alcachofa, espárrago, judías, brécol, coles de Bruselas, repollo, coliflor, col, col rizada, zanahoria, pepino, nabo, berenjena, lechuga, perejil, chirivía, guisante, alfalfa, pimienta, rábano, espinaca, berro, sandía, melón, maíz, sorgo, mijo, trigo, cebada, centeno, avena, 10 arroz, soja, colza, cártamo, maní, batata, yuca, café, coco, piña, cacao, té, plátano, aguacate, higo, guayaba, mango, papaya, olivo, algodón, coníferas, azalea, hortensia, hibisco, rosas, tulipanes, narcisos, petunias, claveles, flor de pascua, y crisantemos.

46. Un método para la producción de un polipéptido de una cualquiera de las reivindicaciones 26 a 33, que comprende:

- 15 a) obtener un organismo recombinante que exprese el polipéptido de la presente invención;
- b) cultivar dicho organismo recombinante;
- c) inducir, de ser necesario, la expresión del polipéptido en dicho organismo recombinante;
- 20 d) purificar el polipéptido a partir del organismo recombinante.

47. Método según la reivindicación 46, en el que el organismo recombinante es una planta transgénica.

25 48. Método según la reivindicación 46, en el que el organismo recombinante es una célula hospedadora de la presente invención, una bacteria recombinante o una levadura recombinante

30 49. Un anticuerpo que se une específicamente a un polipéptido de una cualquiera de las reivindicaciones 26 a 33.

50. Anticuerpo según la reivindicación 49, que se une específicamente al fragmento del polipéptido cuya secuencia:

- a) es la representada por SEQ ID NO:2;
- b) es la secuencia de una delta-endotoxina que está codificada por la secuencia representada por SEQ ID NO:1 o por una secuencia que presenta al menos un 90% de identidad con la misma, o
- 5 c) es la secuencia de una delta-endotoxina que presenta al menos un 80% de identidad con la secuencia representada por SEQ ID NO:2.

51. Anticuerpo según la reivindicación 50, que se une específicamente a la secuencia representada por SEQ ID NO:2.

10

52. Anticuerpo según una cualquiera de las reivindicaciones 49 a 51, que es un anticuerpo monoclonal.

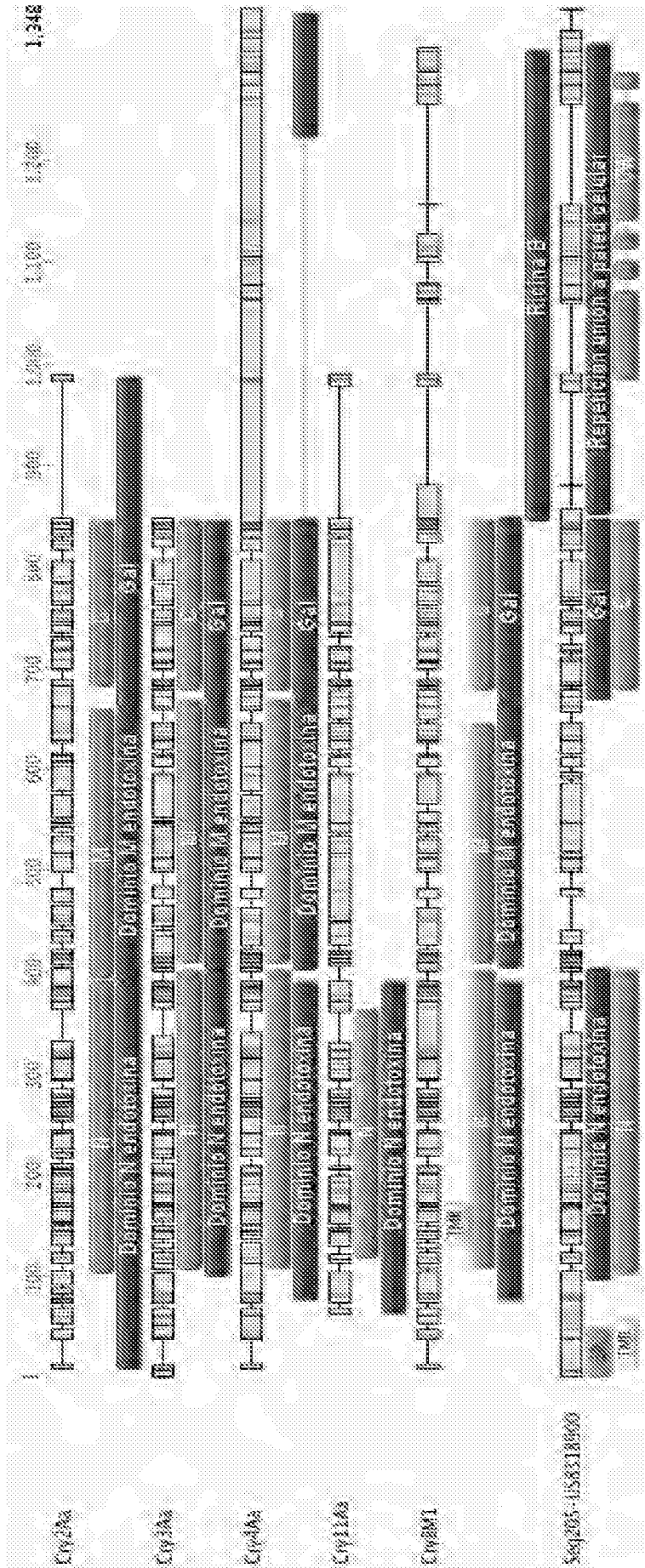


Fig. 1

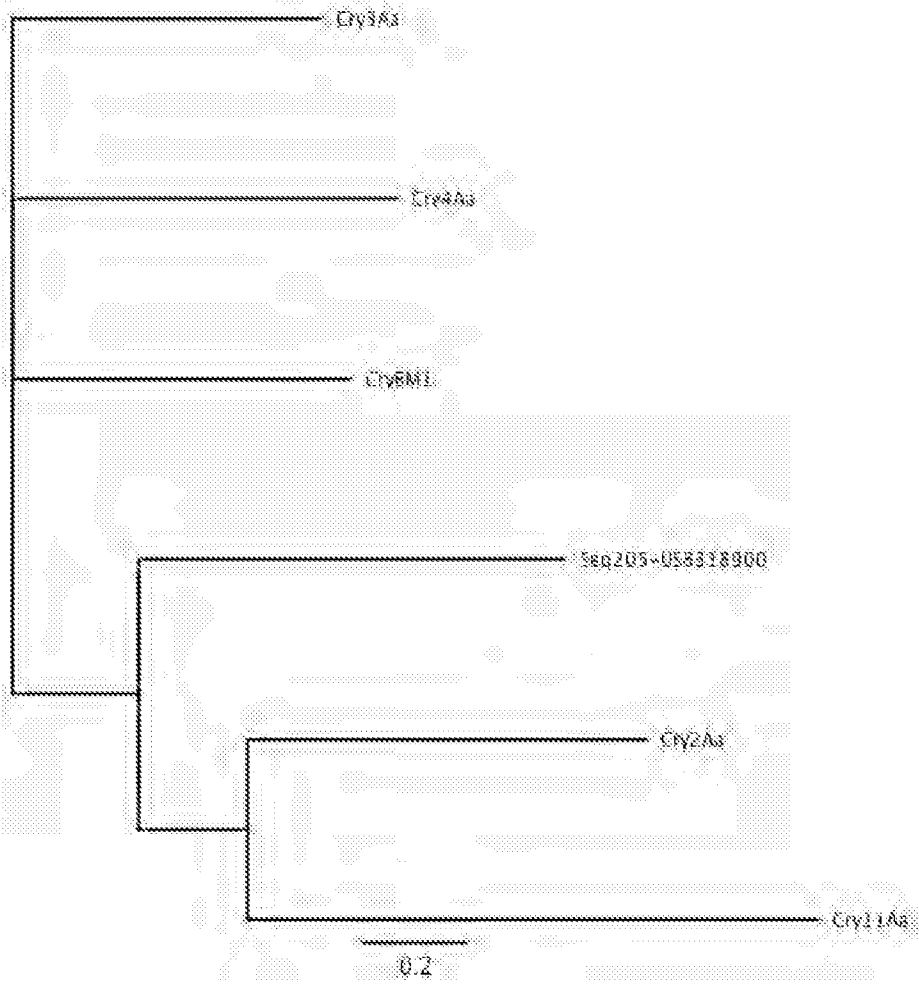


Fig. 2

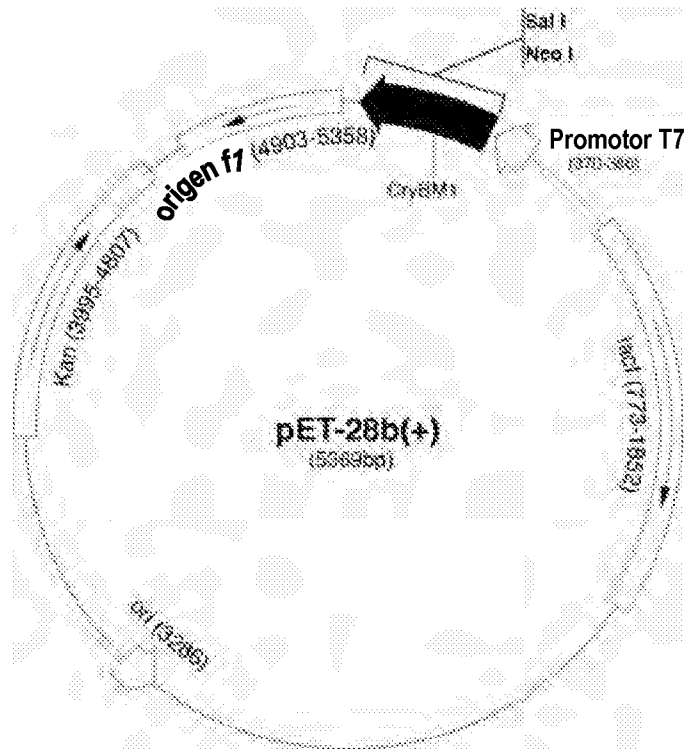


Fig. 3

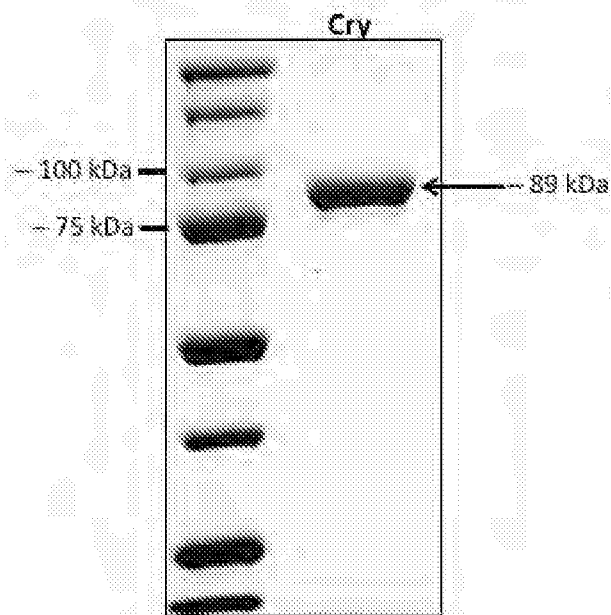


Fig. 4



Fig. 5

ES 2 545 683 B1

LISTADO DE SECUENCIAS

<110> UNIVERSIDAD PUBLICA DE NAVARRA
 Consejo Superior de Investigaciones Científicas

<120> Nueva proteína Cry de Bacillus thuringiensis con actividad
 insecticida para combatir un hemíptero

<130> P-101258

<160> 6

<170> BiSSAP 1.3

<210> 1
 <211> 2397
 <212> DNA
 <213> Bacillus thuringiensis

<220>
 <223> Secuencia codificante de la proteína CryBM1

<220>
 <221> CDS
 <222> 1..2397
 <223> /nota="Proteína CryBM1"
 /tabla_traducción=11

<400> 1
 atg aac caa aat tat aac aac aat gag tat gaa att tta gat aat ggt 48
 Met Asn Gln Asn Tyr Asn Asn Asn Glu Tyr Glu Ile Leu Asp Asn Gly
 1 5 10 15

aat agg agc tat cag act aag tat ccc ttt gca caa gca cca ggt tct 96
 Asn Arg Ser Tyr Gln Thr Lys Tyr Pro Phe Ala Gln Ala Pro Gly Ser
 20 25 30

gaa tgt caa aag ccg aat gcc aaa gat ggg ata tac aca cct gta ggt 144
 Glu Cys Gln Lys Pro Asn Ala Lys Asp Gly Ile Tyr Thr Pro Val Gly
 35 40 45

agg gtg gat aca gtc ctc caa aat ata gat ata ggg tta tct gtt cgt 192
 Arg Val Asp Thr Val Leu Gln Asn Ile Asp Ile Gly Leu Ser Val Arg
 50 55 60

aca gcg ctt tct atc cta caa atg ctt ctg tct gta agt ttt cca gca 240
 Thr Ala Leu Ser Ile Leu Gln Met Leu Leu Ser Val Ser Phe Pro Ala
 65 70 75 80

cta gga aga gct gct gga cta ata aat att att ttt gga ttt ctg tgg 288
 Leu Gly Arg Ala Ala Gly Leu Ile Asn Ile Ile Phe Gly Phe Leu Trp
 85 90 95

gga aca ctt gct ggt caa tct gtt tgg gaa agg ttt atg aaa gcg gta 336
 Gly Thr Leu Ala Gly Gln Ser Val Trp Glu Arg Phe Met Lys Ala Val
 100 105 110

gag tct ctt gta aat caa aaa att aca gat gct gtc aga gca aaa gcc 384
 Glu Ser Leu Val Asn Gln Lys Ile Thr Asp Ala Val Arg Ala Lys Ala
 115 120 125

ES 2 545 683 B1

att	tca	gaa	tta	gaa	ggt	gta	caa	aac	gct	tta	gag	tta	tat	caa	gaa	432
Ile	Ser	Glu	Leu	Glu	Gly	Val	Gln	Asn	Ala	Leu	Glu	Leu	Tyr	Gln	Glu	
	130					135					140					
gca	gca	gat	gat	tgg	aat	gaa	aat	cca	gat	gat	gca	tct	aat	aaa	gag	480
Ala	Ala	Asp	Asp	Trp	Asn	Glu	Asn	Pro	Asp	Asp	Ala	Ser	Asn	Lys	Glu	
145					150					155					160	
cgt	gtt	aga	aga	cag	ttt	aca	tct	acg	aat	aca	ata	atc	gag	tat	gct	528
Arg	Val	Arg	Arg	Gln	Phe	Thr	Ser	Thr	Asn	Thr	Ile	Ile	Glu	Tyr	Ala	
				165					170					175		
atg	cca	tca	ttt	cga	gtt	cca	act	ttt	gaa	gta	ccg	ttg	tta	act	gtg	576
Met	Pro	Ser	Phe	Arg	Val	Pro	Thr	Phe	Glu	Val	Pro	Leu	Leu	Thr	Val	
			180					185					190			
tgt	gcg	caa	gcg	gct	aat	tta	cat	tta	caa	tta	tta	cga	gat	gcc	gtt	624
Cys	Ala	Gln	Ala	Ala	Asn	Leu	His	Leu	Gln	Leu	Leu	Arg	Asp	Ala	Val	
		195					200					205				
aag	ttt	gga	aac	gaa	tgg	ggt	atg	cca	tct	gag	gaa	gta	gag	gat	cta	672
Lys	Phe	Gly	Asn	Glu	Trp	Gly	Met	Pro	Ser	Glu	Glu	Val	Glu	Asp	Leu	
	210					215					220					
tat	act	aga	tta	aca	agg	cgt	acg	gca	gag	tat	aca	gat	tat	tgt	gtg	720
Tyr	Thr	Arg	Leu	Thr	Arg	Arg	Thr	Ala	Glu	Tyr	Thr	Asp	Tyr	Cys	Val	
225					230					235					240	
act	acc	tat	gat	aaa	ggt	tta	aag	ggt	gct	tac	aac	tta	gca	ccg	aat	768
Thr	Thr	Tyr	Asp	Lys	Gly	Leu	Lys	Gly	Ala	Tyr	Asn	Leu	Ala	Pro	Asn	
				245					250					255		
cca	aca	gat	tac	aac	aaa	tac	cct	tat	tta	aat	cca	tat	tcg	aga	gat	816
Pro	Thr	Asp	Tyr	Asn	Lys	Tyr	Pro	Tyr	Leu	Asn	Pro	Tyr	Ser	Arg	Asp	
			260					265					270			
cca	att	tac	ggt	aaa	tac	tat	act	gca	ccc	gtt	gat	tgg	aat	cta	ttt	864
Pro	Ile	Tyr	Gly	Lys	Tyr	Tyr	Thr	Ala	Pro	Val	Asp	Trp	Asn	Leu	Phe	
		275					280					285				
aat	gac	ttt	cgg	aga	gac	atg	aca	att	atg	gta	cta	gac	att	gta	gca	912
Asn	Asp	Phe	Arg	Arg	Asp	Met	Thr	Ile	Met	Val	Leu	Asp	Ile	Val	Ala	
	290					295					300					
gta	tgg	cca	acc	tat	aat	ccg	aga	atc	tat	aca	aat	cca	tat	ggt	gta	960
Val	Trp	Pro	Thr	Tyr	Asn	Pro	Arg	Ile	Tyr	Thr	Asn	Pro	Tyr	Gly	Val	
305					310					315					320	
caa	gta	gaa	ctc	tca	aga	gaa	gta	tat	agc	aca	gta	tat	gga	aga	ggg	1008
Gln	Val	Glu	Leu	Ser	Arg	Glu	Val	Tyr	Ser	Thr	Val	Tyr	Gly	Arg	Gly	
				325					330					335		
ggt	tct	gat	aat	tca	tct	gtg	gag	gca	atc	gaa	tca	caa	att	gtc	aga	1056
Gly	Ser	Asp	Asn	Ser	Ser	Val	Glu	Ala	Ile	Glu	Ser	Gln	Ile	Val	Arg	
			340					345					350			
cca	ccc	cat	tta	gtg	aca	gaa	cta	acc	act	ctt	aaa	att	gag	caa	ggt	1104
Pro	Pro	His	Leu	Val	Thr	Glu	Leu	Thr	Thr	Leu	Lys	Ile	Glu	Gln	Gly	
		355					360					365				

ES 2 545 683 B1

gct act tta gat atg gaa cag ata caa tat ccc aag tat atg aag gtt 1152
Ala Thr Leu Asp Met Glu Gln Ile Gln Tyr Pro Lys Tyr Met Lys Val
370 375 380

acg aat act ctt cat tat att gga agt tct agt aca tgg gag caa agc 1200
Thr Asn Thr Leu His Tyr Ile Gly Ser Ser Ser Thr Trp Glu Gln Ser
385 390 395 400

tct tct gct att cct att aga cca atc acg caa att cat act atc cca 1248
Ser Ser Ala Ile Pro Ile Arg Pro Ile Thr Gln Ile His Thr Ile Pro
405 410 415

gcc aat aat ata ggt aat ctt tca tta tct caa tta gaa gta cct tac 1296
Ala Asn Asn Ile Gly Asn Leu Ser Leu Ser Gln Leu Glu Val Pro Tyr
420 425 430

cga ttt tct ttc tac aac caa gat aat gct tta att gct gaa gtt gga 1344
Arg Phe Ser Phe Tyr Asn Gln Asp Asn Ala Leu Ile Ala Glu Val Gly
435 440 445

tct gaa tcc cct cca aac aat gta acc tgg aac ggg ata cct aga gct 1392
Ser Glu Ser Pro Pro Asn Asn Val Thr Trp Asn Gly Ile Pro Arg Ala
450 455 460

gaa gat agt gat caa aat agt cat cat tta tct tat gtt ggt gct tta 1440
Glu Asp Ser Asp Gln Asn Ser His His Leu Ser Tyr Val Gly Ala Leu
465 470 475 480

agt aca caa tct tca gct ggg ttt cct ggg aca tat cca acg gag ctt 1488
Ser Thr Gln Ser Ser Ala Gly Phe Pro Gly Thr Tyr Pro Thr Glu Leu
485 490 495

ttg ggg gaa tgg gga ttt ggt tgg ttg cat aat agt tta aca ccc gcc 1536
Leu Gly Glu Trp Gly Phe Gly Trp Leu His Asn Ser Leu Thr Pro Ala
500 505 510

aat aga ata gac cca agt aaa atc act caa att cca gcg gtg aaa ggt 1584
Asn Arg Ile Asp Pro Ser Lys Ile Thr Gln Ile Pro Ala Val Lys Gly
515 520 525

ttt gga tta gga gga agt gct aca gtt atc tca gga cct ggt agc aca 1632
Phe Gly Leu Gly Gly Ser Ala Thr Val Ile Ser Gly Pro Gly Ser Thr
530 535 540

ggt ggg gat gta gtg caa tta cag cgg tgg gga cat gta aag ata gca 1680
Gly Gly Asp Val Val Gln Leu Gln Arg Trp Gly His Val Lys Ile Ala
545 550 555 560

att cca gct gca agg aat att cag cct tat cgt gtg aga att cgt tat 1728
Ile Pro Ala Ala Arg Asn Ile Gln Pro Tyr Arg Val Arg Ile Arg Tyr
565 570 575

gca agt agt caa gct agt aca tta aga gta atc cgg tgg agg ggc ggt 1776
Ala Ser Ser Gln Ala Ser Thr Leu Arg Val Ile Arg Trp Arg Gly Gly
580 585 590

gta tat caa cat gca tat tat aat gtt cct caa aca tat acg aat agt 1824
Val Tyr Gln His Ala Tyr Tyr Asn Val Pro Gln Thr Tyr Thr Asn Ser
595 600 605

ES 2 545 683 B1

tta tta aca tat aat aca ttt aaa tat gta gac tcg ttt gct ata aca 1872
 Leu Leu Thr Tyr Asn Thr Phe Lys Tyr Val Asp Ser Phe Ala Ile Thr
 610 615 620

cca gat acc tca atc gag gtt tgg ctt gaa aat acc ggt ggc gga acc 1920
 Pro Asp Thr Ser Ile Glu Val Trp Leu Glu Asn Thr Gly Gly Gly Thr
 625 630 635 640

att atc att gac aaa atc gag ttc att cca aca aca cca aca cca gaa 1968
 Ile Ile Ile Asp Lys Ile Glu Phe Ile Pro Thr Thr Pro Thr Pro Glu
 645 650 655

cca gta gtc gat ggt atc tat caa atc gtg aca gct tta aat aat agt 2016
 Pro Val Val Asp Gly Ile Tyr Gln Ile Val Thr Ala Leu Asn Asn Ser
 660 665 670

agt gtt gca gaa aat ggc gga cca aca aca cgg ggt gga ccg gac cag 2064
 Ser Val Ala Glu Asn Gly Gly Pro Thr Thr Arg Gly Gly Pro Asp Gln
 675 680 685

gtt aaa tta agt cca ttt tac aat tct act gac caa aaa tgg gag ttc 2112
 Val Lys Leu Ser Pro Phe Tyr Asn Ser Thr Asp Gln Lys Trp Glu Phe
 690 695 700

ata tat gat tca aat gaa gat gta tat aca att aga aat tta gct gga 2160
 Ile Tyr Asp Ser Asn Glu Asp Val Tyr Thr Ile Arg Asn Leu Ala Gly
 705 710 715 720

gga ttt ttg act tac ttt atg ttg aac cca gga tat ccg gta tta gca 2208
 Gly Phe Leu Thr Tyr Phe Met Leu Asn Pro Gly Tyr Pro Val Leu Ala
 725 730 735

att cga cct caa tgg gct act gaa aat cag aaa tgg atc gtt gaa ccc 2256
 Ile Arg Pro Gln Trp Ala Thr Glu Asn Gln Lys Trp Ile Val Glu Pro
 740 745 750

gca gga aat ggg tat tat tat ttg aga tct aag tct gtt ccg acc gaa 2304
 Ala Gly Asn Gly Tyr Tyr Tyr Leu Arg Ser Lys Ser Val Pro Thr Glu
 755 760 765

gca gca ttt gca cca aat gct gct gac gga agt gtt gtg aga atg tcg 2352
 Ala Ala Phe Ala Pro Asn Ala Ala Asp Gly Ser Val Val Arg Met Ser
 770 775 780

gct gtt gac ttc agc aca aat caa aaa ttc aaa ctg aat aag tta 2397
 Ala Val Asp Phe Ser Thr Asn Gln Lys Phe Lys Leu Asn Lys Leu
 785 790 795

<210> 2
 <211> 799
 <212> PRT
 <213> Bacillus thuringiensis

<220>
 <223> [CDS]:1..2397 de SEQ ID NO 1

<220>
 <221> DOMINIO

ES 2 545 683 B1

<222> 652..799

<223> Dominio tipo ricina b

<400> 2

Met Asn Gln Asn Tyr Asn Asn Asn Glu Tyr Glu Ile Leu Asp Asn Gly
1 5 10 15
Asn Arg Ser Tyr Gln Thr Lys Tyr Pro Phe Ala Gln Ala Pro Gly Ser
20 25 30
Glu Cys Gln Lys Pro Asn Ala Lys Asp Gly Ile Tyr Thr Pro Val Gly
35 40 45
Arg Val Asp Thr Val Leu Gln Asn Ile Asp Ile Gly Leu Ser Val Arg
50 55 60
Thr Ala Leu Ser Ile Leu Gln Met Leu Leu Ser Val Ser Phe Pro Ala
65 70 75 80
Leu Gly Arg Ala Ala Gly Leu Ile Asn Ile Ile Phe Gly Phe Leu Trp
85 90 95
Gly Thr Leu Ala Gly Gln Ser Val Trp Glu Arg Phe Met Lys Ala Val
100 105 110
Glu Ser Leu Val Asn Gln Lys Ile Thr Asp Ala Val Arg Ala Lys Ala
115 120 125
Ile Ser Glu Leu Glu Gly Val Gln Asn Ala Leu Glu Leu Tyr Gln Glu
130 135 140
Ala Ala Asp Asp Trp Asn Glu Asn Pro Asp Asp Ala Ser Asn Lys Glu
145 150 155 160
Arg Val Arg Arg Gln Phe Thr Ser Thr Asn Thr Ile Ile Glu Tyr Ala
165 170 175
Met Pro Ser Phe Arg Val Pro Thr Phe Glu Val Pro Leu Leu Thr Val
180 185 190
Cys Ala Gln Ala Ala Asn Leu His Leu Gln Leu Leu Arg Asp Ala Val
195 200 205
Lys Phe Gly Asn Glu Trp Gly Met Pro Ser Glu Glu Val Glu Asp Leu
210 215 220
Tyr Thr Arg Leu Thr Arg Arg Thr Ala Glu Tyr Thr Asp Tyr Cys Val
225 230 235 240
Thr Thr Tyr Asp Lys Gly Leu Lys Gly Ala Tyr Asn Leu Ala Pro Asn
245 250 255
Pro Thr Asp Tyr Asn Lys Tyr Pro Tyr Leu Asn Pro Tyr Ser Arg Asp
260 265 270
Pro Ile Tyr Gly Lys Tyr Tyr Thr Ala Pro Val Asp Trp Asn Leu Phe
275 280 285
Asn Asp Phe Arg Arg Asp Met Thr Ile Met Val Leu Asp Ile Val Ala
290 295 300
Val Trp Pro Thr Tyr Asn Pro Arg Ile Tyr Thr Asn Pro Tyr Gly Val
305 310 315 320
Gln Val Glu Leu Ser Arg Glu Val Tyr Ser Thr Val Tyr Gly Arg Gly
325 330 335
Gly Ser Asp Asn Ser Ser Val Glu Ala Ile Glu Ser Gln Ile Val Arg
340 345 350
Pro Pro His Leu Val Thr Glu Leu Thr Thr Leu Lys Ile Glu Gln Gly
355 360 365
Ala Thr Leu Asp Met Glu Gln Ile Gln Tyr Pro Lys Tyr Met Lys Val
370 375 380
Thr Asn Thr Leu His Tyr Ile Gly Ser Ser Ser Thr Trp Glu Gln Ser
385 390 395 400
Ser Ser Ala Ile Pro Ile Arg Pro Ile Thr Gln Ile His Thr Ile Pro
405 410 415
Ala Asn Asn Ile Gly Asn Leu Ser Leu Ser Gln Leu Glu Val Pro Tyr
420 425 430
Arg Phe Ser Phe Tyr Asn Gln Asp Asn Ala Leu Ile Ala Glu Val Gly
435 440 445

ES 2 545 683 B1

Ser Glu Ser Pro Pro Asn Asn Val Thr Trp Asn Gly Ile Pro Arg Ala
 450 455 460
 Glu Asp Ser Asp Gln Asn Ser His His Leu Ser Tyr Val Gly Ala Leu
 465 470 475 480
 Ser Thr Gln Ser Ser Ala Gly Phe Pro Gly Thr Tyr Pro Thr Glu Leu
 485 490 495
 Leu Gly Glu Trp Gly Phe Gly Trp Leu His Asn Ser Leu Thr Pro Ala
 500 505 510
 Asn Arg Ile Asp Pro Ser Lys Ile Thr Gln Ile Pro Ala Val Lys Gly
 515 520 525
 Phe Gly Leu Gly Gly Ser Ala Thr Val Ile Ser Gly Pro Gly Ser Thr
 530 535 540
 Gly Gly Asp Val Val Gln Leu Gln Arg Trp Gly His Val Lys Ile Ala
 545 550 555 560
 Ile Pro Ala Ala Arg Asn Ile Gln Pro Tyr Arg Val Arg Ile Arg Tyr
 565 570 575
 Ala Ser Ser Gln Ala Ser Thr Leu Arg Val Ile Arg Trp Arg Gly Gly
 580 585 590
 Val Tyr Gln His Ala Tyr Tyr Asn Val Pro Gln Thr Tyr Thr Asn Ser
 595 600 605
 Leu Leu Thr Tyr Asn Thr Phe Lys Tyr Val Asp Ser Phe Ala Ile Thr
 610 615 620
 Pro Asp Thr Ser Ile Glu Val Trp Leu Glu Asn Thr Gly Gly Gly Thr
 625 630 635 640
 Ile Ile Ile Asp Lys Ile Glu Phe Ile Pro Thr Thr Pro Thr Pro Glu
 645 650 655
 Pro Val Val Asp Gly Ile Tyr Gln Ile Val Thr Ala Leu Asn Asn Ser
 660 665 670
 Ser Val Ala Glu Asn Gly Gly Pro Thr Thr Arg Gly Gly Pro Asp Gln
 675 680 685
 Val Lys Leu Ser Pro Phe Tyr Asn Ser Thr Asp Gln Lys Trp Glu Phe
 690 695 700
 Ile Tyr Asp Ser Asn Glu Asp Val Tyr Thr Ile Arg Asn Leu Ala Gly
 705 710 715 720
 Gly Phe Leu Thr Tyr Phe Met Leu Asn Pro Gly Tyr Pro Val Leu Ala
 725 730 735
 Ile Arg Pro Gln Trp Ala Thr Glu Asn Gln Lys Trp Ile Val Glu Pro
 740 745 750
 Ala Gly Asn Gly Tyr Tyr Tyr Leu Arg Ser Lys Ser Val Pro Thr Glu
 755 760 765
 Ala Ala Phe Ala Pro Asn Ala Ala Asp Gly Ser Val Val Arg Met Ser
 770 775 780
 Ala Val Asp Phe Ser Thr Asn Gln Lys Phe Lys Leu Asn Lys Leu
 785 790 795

<210> 3
 <211> 25
 <212> DNA
 <213> Secuencia Artificial

<220>
 <223> Cebador directo para amplificar el gen cryBM1

<400> 3
 atgaacccaaa attataacaa caatg

25

<210> 4
 <211> 29

ES 2 545 683 B1

<212> DNA
<213> Secuencia Artificial

<220>
<223> Cebador inverso para amplificar el gen cryBM1

<400> 4
caaatcaaaa attcaaactg aataagtta 29

<210> 5
<211> 31
<212> DNA
<213> Secuencia Artificial

<220>
<223> Cebador directo que añade diana de NcoI

<220>
<221> característica_miscelánea
<222> 1..6
<223> /nota="Diana de NcoI"

<400> 5
ccatggatga accaaaatta taacaacaat g 31

<210> 6
<211> 35
<212> DNA
<213> Secuencia Artificial

<220>
<223> Cebador reverso que añade diana de SalI

<220>
<221> característica_miscelánea
<222> 1..6
<223> /nota="Diana de SalI"

<400> 6
gtcgcactaac ttattcagtt tgaatTTTTg atttg 35