

**CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS
DEL INSTITUTO POLITÉCNICO NACIONAL**

UNIDAD IRAPUATO

**Detección y caracterización de genes *cry* en bacterias
diferentes a *Bacillus thuringiensis***

Tesis que presenta:
M.C. José Francisco Castillo Esparza

Para obtener el grado de:
Doctorado en Ciencias

En la especialidad de:
Biología de Plantas

Director de tesis:
Dr. Jorge Eugenio Ibarra Rendón

El presente trabajo se realizó bajo la dirección del Dr. Jorge Eugenio Ibarra Rendón en el Laboratorio de Bioinsecticidas del Departamento de Biotecnología y Bioquímica, en la Unidad Irapuato del Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional.

Sinodales:

Dra. Gabriela Olmedo Álvarez
Dra. María Cristina del Rincón Castro
Dr. Reynaldo Ariel Álvarez Morales
Dr. Juan José Peña Cabriaes

DEDICATORIA

A mis padres que siempre me han apoyado en todo momento y en todas las decisiones que he tomado a lo largo de mi vida, gracias por su apoyo incondicional, por su comprensión, por sus enseñanzas y por todo el cariño que siempre me han brindado. Los quiero mucho y este gran logro no hubiera sido posible sin ustedes.

A mi hermano que siempre me ha brindado todo su apoyo incondicional para que pudiera culminar esta etapa más en mi vida, gracias por todos los consejos, sé que siempre podré contar contigo para cualquier cosa. Te quiero.

AGRADECIMIENTOS

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo (380133) económico brindado durante 4 años.

Al Centro de Investigación y de Estudios Avanzados del IPN Unidad Irapuato por darme la oportunidad de ser parte de este reconocido Centro de Investigación y permitirme realizar mis estudios de Doctorado.

A mi Director de Tesis, el Dr. Jorge E. Ibarra Rendón por darme la oportunidad de formar parte de su grupo de trabajo, por su confianza y aceptarme en su laboratorio a mitad de semestre, por su paciencia, su asesoría, sus observaciones, su apoyo, amabilidad y disposición en todo momento durante el desarrollo de mi proyecto de tesis, gracias por todo.

A todos los miembros de mi comité tutorial: Dra. Gabriela Olmedo, Dra. María Cristina Del Rincón, Dr. Víctor Olalde y Dr. Ariel Álvarez, por todas sus valiosas aportaciones al proyecto, por todas las críticas constructivas que llevaron a un término satisfactorio del proyecto.

Al Dr. Víctor Olalde por todo el apoyo que me ha brindado desde que ingresé a la maestría y durante estos cuatro años del doctorado. Gracias por sus comentarios, sus opiniones, intervenciones, por las charlas fuera del comité tutorial y sobre todo por su apoyo y amabilidad en todo momento.

Al Dr. Ariel Álvarez por su interés en la realización de este trabajo, por sus sugerencias, comentarios y aportaciones tan acertadas para culminar el proyecto en el tiempo establecido, pero sobre todo por su amabilidad y disponibilidad en todo momento.

A la Dra. Gabriela Olmedo por facilitarnos algunas cepas de su Laboratorio, por todo el apoyo y aportaciones brindadas en la realización del proyecto, por sus sugerencias, comentarios, comprensión y amabilidad y por todas las críticas constructivas que llevaron a un término satisfactorio del proyecto.

A la Dra. Ma. Cristina Del Rincón por la disponibilidad en todo momento, su apoyo, comentarios, sugerencias y aportaciones en el desarrollo del proyecto, por su interés en culminar el proyecto en el tiempo establecido y sobre todo por su amabilidad y comprensión.

Al Dr. Juan José Peña Cabriales por formar parte del comité tutorial, por sus valiosos comentarios, por disposición, interés y amabilidad.

Al Ing. M.C. Javier Luévano Borroel por toda su ayuda en la caracterización de la cepa LBIC-004 y los Bioensayos que se realizaron durante mi tesis de doctorado.

Al M.C. Leandro Gabriel Ordoñez Acevedo por todo el apoyo técnico brindado en diversas técnicas moleculares durante el desarrollo de la tesis.

A la M.C. Africa Islas Robles por el apoyo técnico brindado en el cultivo de cepas y elaboración de medio Marino.

A Ismael Hernández por su valiosa ayuda en la parte bioinformática de mi proyecto.

A Regis por sus enseñanzas, consejos y por recibirme en el laboratorio con una gran calidez humana y hacerme sentir miembro del laboratorio desde el primer día.

A mis compañeros del Laboratorio de Bioinsecticidas: Kathy, Javier, Regis, Pris, Maggie, Nadin, Yari, Moni, Sebas, Gabriel, Mariana, Rosy, Luis, Juventino, Eliana,

Fernando y Pili por su apoyo y por hacer una grata y divertida estancia de cuatro años el laboratorio.

A Kathy, Pris y Maggie por brindarme siempre a tiempo y con amabilidad todo el material que requería en el laboratorio.

A todo el personal de la Biblioteca, Secretaría Académica, Administración, Intendencia, Mantenimiento y Almacén que labora en el CINVESTAV-Unidad Irapuato que hacen que el centro sea funcional y por todas las facilidades prestadas.

A mis amigos, por haber compartido éxitos, viajes, fiestas, frustraciones, enojos, alegrías y fracasos, por su ayuda, cariño, comprensión, consejos y aliento aún en los momentos más difíciles, mil gracias: Damaris, Chio, Crystabel, María, Xóchitl, Javier, Kathy, Margarita, Nancy, Luz, Denis, Gera, Marco, Anny, Moni, Lina, Sebas, Nadin, L. Jorge, Citali, Nelly, Alex, Anacita Pistolas, Vlad, Chino, Medina, Cecilio, Octavio, Mariana, Chente, Yari, Marlene, Daniel, Dalia y “el Pana”.

A mi familia, por el cariño y amor que siempre me han brindado, así como su apoyo incondicional en todos los objetivos y metas que me he planteo.

ÍNDICE GENERAL

	PÁG.
DEDICATORIA	II
AGRADECIMIENTOS	III
ÍNDICE GENERAL	VI
ÍNDICE DE FIGURAS Y TABLAS	IX
RESUMEN	1
ABSTRACT	2
1. INTRODUCCIÓN	3
2. ANTECEDENTES	6
2.1 Evolución del grupo <i>cereus</i>	6
2.2 <i>Bacillus cereus</i>	8
2.3 <i>Bacillus anthracis</i>	8
2.4 <i>Bacillus thuringiensis</i> : Origen e historia	9
2.4.1 Ciclo de vida de <i>Bacillus thuringiensis</i>	10
2.4.2 Características generales.	11
2.4.3 Estructura y clasificación de los genes <i>cry</i> .	12
2.4.4 Modo de acción de las proteínas Cry	16
2.4.5 Ecología de <i>Bacillus thuringiensis</i>	18
2.5 Identificación de genes <i>cry</i> en bacterias diferentes a <i>B. thuringiensis</i>	21
2.6 Sistema de iniciadores universales	23
2.7 CD-HIT y su Algoritmo de agrupamiento	24
2.8 HMMER	25
2.8.1 Aplicaciones de perfiles HMM	26
2.8.2 Trabajos relevantes con el uso de perfiles HMM	28
3. JUSTIFICACIÓN	32
4. HIPÓTESIS	32
5. OBJETIVOS	33
5.1 General	33

5.2	Específicos	33
6.	MATERIALES Y MÉTODOS	34
6.1	Búsqueda de genes <i>cry</i> en cepas de <i>Bacillus</i> spp.	34
6.1.1	Selección de cepas y condiciones de crecimiento	34
6.1.2	Extracción de DNA genómico	35
6.1.3	Amplificación de los posibles genes <i>cry</i>	35
6.1.4	Clonación, transformación, secuenciación e identificación de los posibles genes <i>cry</i>	36
6.1.5	Obtención e identificación de la secuencia completa del gen <i>cry</i>	36
6.1.6	Transformación de las cepas acristalíferas de <i>Bt</i>	38
6.1.7	Análisis de proteínas	38
6.1.8	Caracterización de la cepa	39
6.1.9	Bioensayos	40
6.2	Búsqueda <i>in silico</i> de proteínas Cry en genomas de <i>Bacillus</i> spp	41
6.2.1	Recuperación de secuencias de las proteínas Cry	41
6.2.2	Agrupación y comparación de secuencias con CD-HIT	41
6.2.3	Construcción del perfil HMM	42
6.2.4	Validación del perfil HMM y análisis de los datos obtenidos por el modelo	43
6.2.5	Análisis MLST y reconstrucción filogenética	44
7.	RESULTADOS	45
7.1	Identificación y caracterización de la cepa LBIC-004	47
7.2	Patrón de plásmidos	48
7.3	Identificación, clonación, análisis y clasificación del gen <i>cry</i>	50
7.4	Expresión y análisis de perfil de proteínas (SDS-PAGE)	52
7.5	Bioensayos	54
7.6	Agrupamiento de proteínas Cry obtenidos mediante el uso de CD-HIT y construcción del perfil HMM	54
7.7	Análisis del plásmido pBtoxis y Cepa 4Q7 con el perfil HMMs	55
7.8	Identificación de proteínas Cry mediante <i>hmmsearch</i> en genomas de <i>Bacillus</i>	57

7.9	Análisis MLST y reconstrucción filogenética	60
8	DISCUSIÓN	65
9	CONCLUSIÓN	76
10	PERSPECTIVAS	78
11	REFERENCIAS	79

ÍNDICE DE FIGURAS Y TABLAS

Figura 1. Bloques de aminoácidos conservados (1-8) entre las diferentes proteínas Cry de tres dominios	14
Figura 2. Estructura tridimensional de la toxina Cry2Aa	15
Figura 3. Modelos del modo de acción de las toxinas Cry y mecanismo resultante para la resistencia	17
Figura 4. Diagrama de Venn de tres conjuntos de cepas que amplifican con cualquiera de los tres sets de iniciadores	45
Figura 5. Alineamiento de un fragmento de la secuencia de nucleótidos de los genes parciales obtenidos de las cepas LBICR6, LBICS9, LBICT4, LBICT7, LBICY4, LBIC-004.	47
Figura 6. Amplificación de los genes <i>16S rRNA</i> , <i>gyrB</i> y <i>hag</i>	48
Figura 7. Patrón de plásmidos de la cepa LBIC-004	49
Figura 8. Amplificación parcial del gen <i>cry-like</i> y amplificación completa y clonación del gen <i>cry-like</i>	51
Figura 9. Estructura terciaria de la proteína Cry-like	52
Figura 10. Estado de esporangio de la cepa nativa y las cepas recombinantes de <i>B. thuringiensis</i> .	52
Figura 11. Perfil proteico de la cepa LBIC-004 y 4Q7 recombinante	53
Figura 12. Diagrama para construcción del perfil HMM para búsqueda de proteínas Cry.	56
Figura 13. Distribución de las proteínas Cry obtenidas del análisis <i>in silico</i> de 857 genomas de <i>Bacillus</i> mediante el perfil HMM.	58
Figura 14. Árbol filogenético del análisis de MLST	64
Tabla 1. Características generales de los iniciadores utilizados para la caracterización de la cepa LBIC-004 y el gen <i>cry-like</i> .	39
Tabla 2. Cepas analizadas que mostraron un porcentaje de identidad con genes <i>cry</i> .	47
Tabla 3. Resultados obtenidos utilizando el perfil HMM.	57

Tabla 4. Principales subgrupos de proteínas Cry identificadas <i>in silico</i> por el modelo HMM en 857 genomas de <i>Bacillus</i> spp	59
Tabla 5. Subgrupos de proteínas Cry encontrados en especies diferentes a <i>B. thuringiensis</i> mediante el modelo HMM.	60
Tabla 6. Identificación por MLST de las bacterias cuyos genomas mostraron secuencias Cry, y su comparación con las especies registradas en el GenBank, usadas para el análisis <i>in silico</i> de este estudio y el perfil alélico encontrado para cada cepa.	61
ANEXO 1. Patrón de plásmidos de las cepas LBIC-R6, LBIC-S9, LBIC-T4, LBIC-T7, LBIC-Y4, LBIC-004 y análisis del gen <i>16S rRNA</i> y <i>gyrB</i> de las cepas LBIC-R6, LBIC-S9, LBIC-T4, LBIC-T7, LBIC-Y4, LBIC-004.	93
ANEXO 2. Análisis de restricción con diferentes enzimas de las clonas obtenidas del amplicon de la cepa LBIC-004.	93
ANEXO 3. Secuencia del nuevo gen <i>cry-like</i> de la cepa LBIC-004.	95
ANEXO 4. Peptidos obtenidos de la proteína de la cepa LBIC-004 y 4Q7 recombinante.	96
ANEXO 5. Bioensayos realizados con diferentes insectos y nemátodos.	97
ANEXO 6. Datos obtenidos de los 857 genomas obtenidos de la base de datos del GenBank y analizados para la búsqueda de proteínas Cry con el modelo HMM.	99
ANEXO 7. Información sobre algunas cepas de <i>Bacillus</i> spp. utilizadas en el proyecto, su origen y características fenotípicas	106
ANEXO 8. Publicación obtenida del trabajo de tesis	111

RESUMEN

Bacillus thuringiensis (*Bt*) es el bioinsecticida más exitoso contra diferentes plagas en la agricultura y en el sector de salud pública. Su actividad se atribuye a las proteínas Cry que expresa en su fase de esporulación, sin embargo, estas proteínas no son exclusivas de *Bt*, ya que hay algunos reportes donde han identificado genes *cry* en otras especies de *Bacillus* y en otros géneros. Por esta razón, se trabajó con una colección de 223 cepas de *Bacillus* en la cual se excluyó a *Bt*. Estas cepas se sometieron a escrutinio para la identificación de genes *cry*, de las cuales sólo 6 cepas mostraron resultados positivos: LBICR6, LBICS9, LBICT4, LBICT7, LBICY4 y LBIC-004. Sin embargo, los genes *cry* identificados en estas cepas mostraron el mismo porcentaje de identidad parcial (65%) con la proteína Cry8Ca1; un alineamiento de las secuencias de cada cepa mostró que son réplicas del mismo aislado. Este aislado no es capaz de formar inclusiones proteicas en la fase de esporulación, además mostró un alto porcentaje de identidad a *B. cereus* utilizando los genes *16S rRNA*, *gyrB* y *hag*. Por otro lado, se realizó un patrón de plásmidos de las 6 cepas para su caracterización, todas mostraron el mismo número y tamaño de plásmidos. En cuanto al gen *cry*, éste está compuesto de 1953 pb y codifica para una proteína de 651 aa (74.9 kDa), para la cual se predijo mediante el software Phyre2 una estructura tridimensional de tres dominios. La secuencia de aminoácidos de la proteína tiene una identidad del 41% con la proteína Cry8Ca, lo cual indica que el gen de la cepa LBIC-004 es un gen *cry* totalmente nuevo (*cry-like*). El gen *cry-like* fue transformado en una cepa acristalífera de *Bt* (4Q7-*cry-like*) la cual fue utilizada, al igual que la cepa LBIC-004, para realizar bioensayos contra *Caenorhabditis elegans* y algunos insectos como *Aedes aegypti*, *Manduca sexta* y *Phyllophaga* sp; ninguna de las cepas mostro toxicidad contra *C. elegans* o con los otros insectos probados. Por otra parte, se realizó la identificación *in silico* de genes *cry* en genomas de *Bacillus* mediante el uso de un perfil de Modelos Ocultos de Markov diseñado a partir del extremo C-terminal de las proteínas Cry. Con este modelo se analizaron 857 genomas de *Bacillus* y se logró identificar 174 proteínas Cry, de las cuales 42 correspondían a *B. cereus*, *B. spp.* y *B. subtilis*, el resto correspondía a *Bt*.

ABSTRACT

Bacillus thuringiensis (*Bt*) is known as the most successful microbial insecticide against different pests in agriculture and in the public health sector. Its activity is attributed to the Cry proteins that it expresses in its sporulation phase. However, these proteins are not unique to *Bt*, as there are some reports where they have identified *cry* genes in other *Bacillus* species and even in other genera. For this reason, we worked with a collection of 223 strains of *Bacillus* in which *Bt* was excluded. These strains were screened for the identification of *cry* genes, only 6 strains tested positive: LBICR6, LBICS9, LBICT4, LBICT7, LBICY4 and LBIC-004. However, these strains showed the same percentage of partial identity (65%) with the Cry8Ca1 protein; an alignment of the sequences of each strain showed that they are replicas of the same isolate. This isolate is not able to form protein inclusions in the sporulation phase, it also showed a high percentage of identity to *B. cereus* using the 16S rRNA, *gyrB* and *hag* genes. On the other hand, a plasmid pattern of the 6 strains was made for characterization, all showed the same number of plasmids and size. As for the *cry* gene, it is 1953 bp long that codes for a 651aa protein with a molecular weight of 74.9kDa, which has a three-dimensional structure of three domains predicted by the Phyre2 software, as well as three-domain Cry proteins. The amino acid sequence of the protein has a 41% identity with the Cry8Ca protein, which indicates that the gene of strain LBIC-004 is a totally new *cry* gene (*cry-like*). The *cry-like* gene was transformed into an acristaliferous strain of *Bt* (4Q7-*cry-like*) which was used as well as strain LBIC-004 to perform bioassays against *Caenorhabditis elegans* and some insects such as *Aedes aegypti*, *Manduca sexta* and *Phyllophaga* sp.; none of the strains showed toxicity against *C. elegans* or with the other tested insects. On the other hand, the *in silico* identification of *cry* genes in *Bacillus* genomes was performed by using a profile of Hidden Markov Models designed from the carboxyl terminal end of the Cry proteins. With this model, 857 *Bacillus* genomes were analyzed and 174 Cry proteins were identified, of which 42 proteins corresponded to *Bacillus cereus*, *Bacillus* spp. and *Bacillus subtilis*, the rest corresponded to *Bt*.

1. INTRODUCCIÓN

Bacillus thuringiensis (*Bt*) es una bacteria aerobia, Gram-positiva, de flagelación períttrica, posee la característica de desarrollar esporas de resistencia de forma elipsoidal que no provocan hinchamiento bacilar (Sauka y Benintende 2008). *Bt* ha sido ampliamente utilizado como bio-insecticida contra plagas de insectos en los sectores de agricultura y salud pública; su uso se debe principalmente a que durante el proceso de esporulación produce una inclusión parasporal de naturaleza proteica formada por proteínas Cry y Cyt, estas proteínas son tóxicas para una gran variedad de insectos como lepidópteros, dípteros, coleópteros y contra algunos nemátodos y protozoos (Reinoso-Pozo, Del Rincón-Castro e Ibarra, 2016). Los genes *cry* tienen una secuencia codificante de entre 1900 y 3600 pares de bases (pb). Las proteínas que codifican presentan una identidad aminoacídica que puede variar en gran medida desde más de un 90% hasta menos de un 20% y el intervalo de masas moleculares oscila entre 50 y 140 kDa (Bravo *et al.*, 2013).

Las proteínas Cry son propiamente protoxinas que se solubilizan en el medio alcalino del intestino de los insectos y luego son proteolíticamente procesadas en un fragmento tóxico de menor tamaño. El extremo amino-terminal (N_t) de la proteína codifica para el dominio tóxico, mientras que el extremo carboxi-terminal (C_t) se cree que está relacionado con la formación y estabilización del cristal (Palma *et al.*, 2014). Las dos partes de las proteínas también difieren en la conservación de las secuencias de aminoácidos. La parte C_t de las proteínas está altamente conservada, mientras que el extremo N_t presenta una mayor variación, concentrándose en regiones hipervariables que determinan la especificidad insecticida (Ge, Shivarova y Dean, 1989; Evdokimov *et al.*, 2014). Hasta la fecha se han clonado y secuenciado más de 800 genes *cry* diferentes (Crickmore *et al.*, 2018). Por otra parte, las proteínas Cry de tres dominios muestran claras diferencias en sus secuencias de aminoácidos, a pesar de esto, todas tienen en

común una estructura terciaria muy similar y conservada, la cual consta de tres dominios (Palma *et al.*, 2014).

Por otro lado, existen algunos autores quienes han encontrado toxinas Cry en otras especies de bacterias diferentes a *Bt*. En el grupo de Barloy (1996) clonaron y expresaron el primer gen tóxico (*cry16Aa1*) de *Clostridium bifermentans* que codifica una nueva proteína mosquitocida con homología a las δ -endotoxinas de *Bt*. En 1998 reportaron la clonación y expresión de un segundo gen en *C. bifermentans* (*cry17Aa1*) que codifica una proteína de 71 kDa sin toxicidad significativa hacia larvas de mosquitos.

Otro grupo de investigación detectó un cristal paraesporal (Cry18Aa1) en extractos proteicos de esporangios de *Paenibacillus popilliae* aislado de una larva enferma de *Melolontha melolontha* L., un análisis de aminoácidos reveló homología significativa a la δ -endotoxina Cry2Aa de *Bt*. (Zhang *et al.*, 1997).

Por otro lado, Yokoyama (2004) seleccionó una clona de *Paenibacillus lentimorbus* que contiene dos genes *cry* (*cry43Aa1*, *cry43Ba1*), un gen *cry* parcial (*cry43-like*), y tres genes más pequeños situados río arriba. La proteína Cry43Aa1 provocó inhibición de la ingestión y un 90% de mortalidad en larvas de primer estadio de *Anomala cuprea*.

Un estudio más reciente mostró la presencia de un tipo de toxina de dos componentes en *Bacillus sphaericus* (*Lysinibacillus sphaericus*) que no había sido descrita. La actividad insecticida de esta proteína, Cry48Aa1, es sólo activada en presencia de una segunda proteína, en este caso Cry49Aa1. Ambos componentes Cry48Aa1 y Cry49Aa1 son necesarios para la patogenicidad, por lo tanto, indica una interacción sin precedentes para generar toxicidad. Esta nueva combinación de toxinas puede indicar una combinación fortuita de los miembros de las familias de genes que codifican toxinas Cry de tres dominios y las toxinas binarias (Jones *et al.*, 2007).

Sin embargo, a pesar del gran número de proteínas Cry reportadas hasta la fecha, es importante señalar que tienen un rango muy estrecho de hospederos y muy pocas cepas tienen alta toxicidad. Por ello, es necesario intensificar la búsqueda de nuevas proteínas Cry con el fin de identificar proteínas más tóxicas, diferentes o nuevas a las ya utilizadas dentro del control biológico, las cuales posean actividad insecticida contra insectos que no han podido ser controlados con *Bt* y constituyan una alternativa para el control de insectos que ya han desarrollado resistencia a dichas toxinas. Además, ya que se han encontrado genes *cry* en bacterias diferentes a *Bt*, es de gran importancia hacer un estudio sistemático de búsqueda de estos genes en otras bacterias, lo cual podría dar información acerca de la distribución de estos genes y sobre todo, ver qué tan específicos son de *Bt*.

En este trabajo se reporta una búsqueda exhaustiva de genes *cry* en bacilos, tanto experimentalmente como *in silico*. Se realizó una búsqueda de genes *cry* mediante PCR en un total de 223 cepas en las cuales se logró amplificar e identificar un gen *cry* totalmente nuevo. Por otra parte, la cepa LBIC-004 mostró un alto porcentaje de identidad a *B. cereus* utilizando los genes *16S rRNA*, *gyrB* y *hag*. Además, se realizó una búsqueda *in silico* de proteínas Cry mediante el desarrollo de un perfil de Modelos Ocultos de Markov diseñado a partir del extremo carboxilo de las proteínas Cry reportadas, mediante este modelo se analizaron 857 genomas de bacilos reportados hasta el 2017 y se lograron identificar 174 proteínas Cry distribuidas dentro del grupo *Bacillus cereus*.

2. ANTECEDENTES

Las bacterias del grupo *Bacillus cereus* incluyen a *Bacillus thuringiensis*, *Bacillus anthracis* y *Bacillus cereus (sensu stricto)*. Este grupo de microorganismos son formadores de esporas, bacterias Gram-positivas que forman una subdivisión altamente homogénea dentro del género *Bacillus*. La demostración de su alta relación genética ha contribuido a la sugerencia de que *B. anthracis*, *B. cereus* y *B. thuringiensis* son miembros de una sola especie, *Bacillus cereus sensu lato*. Tradicionalmente, estos organismos se han diferenciado en función de sus características fenotípicas, incluido el potencial patogénico. *Bacillus mycoides*, *Bacillus pseudomycoides* y *Bacillus weihenstephanensis* también son miembros del grupo *cereus* (Chen y Tsen 2002).

2.1 Evolución del grupo *cereus*

El grupo *B. cereus* comprende bacterias estrechamente relacionadas que exhiben propiedades patógenas altamente divergentes. Muchas bacterias clasificadas como *B. cereus* están ampliamente distribuidas en el medio ambiente, con reservorios probables en el suelo y como habitantes comensales de los intestinos de los insectos (Margulis *et al.* 1998). Ocasionalmente se asocian con intoxicación alimentaria y con infecciones de tejidos blandos, particularmente del ojo. Otros miembros del grupo que actualmente están clasificados como *Bacillus thuringiensis* son principalmente patógenos de insectos. Estas bacterias producen toxinas en forma de proteínas cristalinas paraesporales que se han usado ampliamente para el control biológico de plagas de insectos (Schnepf, *et al.* 1998). Ocasionalmente, las cepas de *B. thuringiensis* son responsables de infecciones humanas similares a las causadas por cepas de *B. cereus*. Un tercer fenotipo patógeno es exhibido por *Bacillus anthracis*, un patógeno de mamíferos y especialmente de ungulados que pueden causar enfermedades en humanos. Los principales factores de virulencia de *B. anthracis* están codificados por genes ubicados en dos plásmidos: los genes de la toxina tripartita *pag*, *lef* y *cya* están

contenidas en el plásmido pXO1, mientras que los genes que codifican la biosíntesis de la cápsula de poli-D-glutamato, *capA*, *capB*, y *capC*, se transportan en un plásmido más pequeño, pXO2 (Okinaka *et al*, 1999). De manera similar, las proteínas cristalinas de los aislados de *B. thuringiensis* son las principales responsables de la toxicidad característica hacia los insectos y casi invariablemente son codificadas por plásmidos. Los genes de virulencia de *B. cereus*, por otro lado, son cromosómicos.

Estas tres especies principales del grupo *B. cereus* se describieron por primera vez a fines del siglo XIX, sin embargo, a pesar de esta larga historia, las relaciones entre estos organismos aún no se han resuelto por completo. La hibridación de DNA del genoma completo no ha sido útil, mientras que los marcadores convencionales de diversidad cromosómica, tales como los genes del 16S *rRNA* y 23S, son esencialmente idénticos. Estudios completos que utilizan una amplia gama de técnicas, que incluyen mapeo genómico, electroforesis en gel de campo pulsado de DNA cromosómico, electroforesis de enzimas multilocus, mapeo repetido en tándem de número variable, toma de huellas dactilares BOX-PCR, el análisis del polimorfismo de longitud de fragmentos amplificados (AFLP) y la tipificación de secuencia multilocus (MLST), han revelado amplias similitudes genómicas y pocas diferencias consistentes entre los aislados actualmente clasificados como *B. anthracis*, *B. cereus* y *B. thuringiensis*. Estos estudios han reforzado el argumento fenotípico de que los tres taxones deberían considerarse una sola especie bacteriana (Priest *et al*. 2004).

A pesar de tales argumentos biológicos para la unificación, se ha mantenido un estado de especie separado para estas bacterias debido a sus características patogénicas distintivas. Prácticamente todos los aislamientos del grupo *cereus* obtenidos de humanos o animales que presentan los síntomas del ántrax están estrechamente relacionados entre sí, y es muy probable que *B. anthracis* sea un clon, particularmente si está asociado con los plásmidos que codifican la toxina pXO1 y pXO2 (Keim *et al*. 2000). Sin embargo, los organismos clasificados como

B. cereus y *B. thuringiensis* son más diversos, y las relaciones evolutivas entre todos los miembros del grupo aún no se han establecido definitivamente (Hill *et al.*, 2004). Esto es importante, no sólo para comprender la evolución de la virulencia en el grupo *cereus*, sino también para caracterizar estos organismos rápidamente y con precisión, una preocupación que ha adquirido una creciente importancia científica y política en los últimos años.

2.2 *Bacillus cereus*

B. cereus es de naturaleza ubicua y un patógeno oportunista, a menudo asociado con dos formas de intoxicación alimentaria humana, que se caracteriza por diarrea y malestar abdominal o náuseas y vómitos. En individuos sanos, pero principalmente en individuos con ciertas afecciones subyacentes, como pacientes inmunocomprometidos o pacientes que se recuperan de una cirugía, se sabe que *B. cereus* causa una variedad de infecciones, que incluyen: endoftalmitis, bacteriemia, septicemia, endocarditis, salpingitis, infecciones cutáneas, neumonía y meningitis (Logan *et al.*, 1999). *B. cereus* se encuentra como contaminante en muchos productos alimenticios, incluidos los productos lácteos. Sin embargo, su nicho ecológico primario es el ambiente del suelo. También se encuentra comúnmente como parte de la microflora intestinal de los invertebrados, no sólo como esporas sino también como células vegetativas en crecimiento. Por definición, *B. cereus* es acristalífero, pero una cepa de *B. cereus* que porta un gen *cry* funcional se considera una cepa de *B. thuringiensis*. No se han identificado factores de virulencia específicos para *B. cereus* y las proteínas que se cree que son específicas de *B. cereus* se han encontrado recientemente en aislados de *B. thuringiensis* (Kolstø *et al.*, 2002).

2.3 *Bacillus anthracis*

B. anthracis es el agente etiológico del ántrax o carbunco, una enfermedad mortal aguda que se encuentra principalmente entre los herbívoros, pero de hecho

todos los mamíferos son susceptibles. En los últimos años, se ha hecho más conocido por su uso como arma biológica. La bacteria es endémica en África, Asia y América del Sur, pero se puede encontrar en gran parte del mundo. Las esporas de *B. anthracis* son altamente resistentes al calor, las radiaciones ultravioletas e ionizantes, la presión y una variedad de agentes químicos y se cree que sobreviven en el medio ambiente durante décadas y posiblemente siglos (Turnbull 2002). A pesar de que *B. anthracis* comparte el mismo nicho ecológico que *B. cereus* y *B. thuringiensis* y puede aislarse fácilmente de muestras de suelo contaminadas, aún no está claro si su ciclo de vida incluye una etapa vegetativa fuera del huésped. Las esporas son ingeridas por los herbívoros y germinan dentro del huésped para producir células vegetativas, que se multiplican y producen factores de virulencia y finalmente matan al huésped. Al morir, se liberan grandes cantidades de bacilos al medio ambiente y esporulan al contacto con el aire completando así el ciclo de vida (Mock y Fouet 2001).

B. anthracis puede diferenciarse de *B. cereus* y *B. thuringiensis* mediante pruebas microbiológicas y bioquímicas. Los aislados de *B. anthracis* no son hemolíticos, no móviles, sensibles a la penicilina, susceptibles a fagos específicos y producen una cápsula de ácido poli- γ -d-glutámico (Hoffmaster *et al.*, 2004). Al igual que *B. thuringiensis*, la capacidad de *B. anthracis* para causar una enfermedad se atribuye principalmente a su contenido plasmídico. Cepas completamente virulentas de *B. anthracis* portan dos plásmidos grandes, pXO1 (181 kb) y pXO2 (96 kb), que codifican la maquinaria necesaria para producir y regular los factores de virulencia del ántrax, la toxina tripartita y la cápsula, respectivamente (Okinaka *et al.*, 1999).

2.4 *Bacillus thuringiensis*: Origen e historia

La era de *Bacillus thuringiensis* tuvo su comienzo en 1901 cuando un científico japonés llamado Shigetane Ishiwata aisló una bacteria a partir de larvas de gusanos de seda muertas mientras él estaba investigando la causa de la

llamada “flacheria” y que él denominó "enfermedad sotto". Esta enfermedad fue responsable de la pérdida de un gran número de gusanos de seda en Japón y la región circundante. Ishiwata nombró a la bacteria como *Bacillus sotto*, sin hacer la descripción formal de una nueva especie. Unos años después Ernst Berliner (1911) aisló una cepa bacteriana, la cual estaba relacionada con las larvas de polilla que encontró muertas en un molino de harina en la provincia alemana de Turingia. Él apropiadamente describió y nombró al microorganismo como *Bacillus thuringiensis*. Berliner estudió la bacteria y encontró unos cuerpos de inclusión o "Restkorpen", junto a la endospora (Berliner, 1911). Mattes en 1927 observó de nuevo los mismos cuerpos de inclusión en *Bt*, pero no fue hasta mucho más tarde (25 años) que la actividad insecticida fue atribuida a estos cuerpos de inclusión refringentes ahora denominados "cristales paraesporales", una frase acuñada por Christopher Hannay en 1953. Thomas Angus junto con Philip Fitz-James y Hannay en 1955, demostraron rápidamente que los cuerpos de inclusión producen una actividad insecticida y descubrieron que éstos se componen de una o varias proteínas (Ibrahim *et al.*, 2010).

2.4.1 Ciclo de vida de *Bacillus thuringiensis*

El ciclo de vida de *B. thuringiensis* es caracterizado por dos fases, la cual incluye la división celular vegetativa y el desarrollo de la espora, otras veces referido como ciclo de esporulación. La célula vegetativa tiene forma de barra y se divide en dos células hijas uniformes mediante la formación de un septo de división iniciado a mitad de la membrana plasmática. La esporulación por otro lado involucra la división celular asimétrica y es caracterizada por siete etapas, los cuales incluyen la formación del filamento axial (etapa I), formación del septo de la preespora (etapa II), englobamiento (etapa III), primera aparición del cristal paraesporal y formación de la preespora (etapas IV a VI), formación del exosporio, pared celular primordial, córtex y capas de esporas acompañadas por la transformación del nucleoide de la espora (etapa VII); maduración de la espora y lisis del esporangio. La producción de proteínas cristalinas por *Bt* durante la

esporulación es un fenómeno biológico único regulado genéticamente, que probablemente alivia el estrés físicamente al compensar la pérdida de agua durante la formación de esporas y proporciona una ventaja adicional de supervivencia al ejercer una acción letal contra los insectos hospederos. A su vez, la acción tóxica proporciona suficientes nutrientes del huésped para permitir la germinación de la espora bacteriana latente y su retorno al crecimiento vegetativo (Ibrahim *et al.*, 2010; Bulla *et al.*, 1980; Bechtel y Bulla 1982).

2.4.2 Características generales.

B. thuringiensis es un bacilo Gram-positivo, de flagelación peritrica, que mide de 3 a 5 μm de largo por 1 a 1.2 μm de ancho y que posee la característica de desarrollar esporas de resistencia elipsoidales que no provocan el hinchamiento del perfil bacilar; se ha aislado de todo el mundo a partir de una gran diversidad de ecosistemas, incluidos el suelo, agua, insectos muertos, polvo, hojas de árboles, diversas coníferas y mamíferos insectívoros, así como de los tejidos humanos con necrosis severa (Sauka y Benintende, 2008).

Es un microorganismo anaerobio facultativo, quimioorganotrofo y con actividad de catalasa; sin embargo, la característica principal de *B. thuringiensis* es que durante el proceso de esporulación produce una inclusión paraesporal formada por uno o más cuerpos cristalinos de naturaleza proteica que son tóxicos para distintos invertebrados, especialmente larvas de insectos. Estas proteínas se llaman δ -endotoxinas y constituyen la base del insecticida biológico más difundido a nivel mundial. Estas δ -endotoxinas son sintetizadas por cepas de *Bt* en el inicio de la esporulación y durante la fase de crecimiento estacionario como inclusiones cristalinas parasporales y pueden ser de dos tipos: Cry (cristal) y Cyt (citolíticas) (Palma *et al.*, 2014). Una vez ingeridos por los insectos, estos cristales son solubilizados gracias al pH alcalino del intestino medio del insecto, las toxinas son activadas proteolíticamente por proteasas del intestino medio y se unen a receptores específicos localizados en la membrana celular del insecto, lo que

conduce a la alteración celular y la muerte del insecto (Sauka y Benintende 2008; Deist *et al.*, 2014; Palma *et al.*, 2014).

2.4.3 Estructura y clasificación de los genes *cry*.

Desde la identificación y clonación del primer gen de una proteína Cry en 1981, el número de genes que codifican para nuevas proteínas insecticidas se ha incrementado significativamente, generando la necesidad de un sistema de nomenclatura organizada. En el primer sistema, los nombres de las toxinas Cry y sus correspondientes genes incluyen un número romano dependiendo de la actividad insecticida de la proteína Cry, es decir, Cry I para proteínas tóxicas para lepidópteros (CryIA, CryIB, CryIC, CryID), Cry II para proteínas con toxicidad contra lepidópteros y dípteros (CryIIA, CryIIB), Cry III para proteínas tóxicas contra coleópteros (CryIIIA) y Cry IV para proteínas tóxicas exclusivamente contra dípteros (CryIVA, CryIVB, CryIVC y CryIVD). Sin embargo, este sistema exhibe complicaciones importantes, por ejemplo, la actividad de nuevas toxinas tuvo que ser ensayada contra una lista creciente de insectos antes de que el gen y la toxina pudiera ser nombrado, algunas proteínas homólogas nuevas eran, de hecho, no tóxicas como se esperaba, y otros exhibieron toxicidad dual frente a especies de dípteros y lepidópteros, o dípteros y coleópteros (Palma *et al.*, 2014; Méndez-López *et al.*, 2003).

Para evitar estos problemas, en 1998 Crickmore *et al.*, propusieron una revisión de la nomenclatura original que ha sido rápidamente aceptada. Estos autores plantean la clasificación de las proteínas cristalinas de *B. thuringiensis* basándose en el sistema de nomenclatura de la súperfamilia del citocromo P-450, dependiendo ésta únicamente de datos de secuencia. Esto permite clasificar juntas a toxinas estrechamente emparentadas, evitándose la necesidad de bioensayar cada nueva proteína antes de asignarle un nombre. La nueva nomenclatura define a una proteína Cry del siguiente modo: una proteína cristalina de *B. thuringiensis* que presenta un efecto tóxico experimentalmente verificable

contra un organismo susceptible, o bien cualquier proteína con una evidente similitud de secuencia con una proteína Cry conocida. Los números romanos han sido sustituidos en la nueva nomenclatura por números arábigos en el primer rango de la clasificación, el uso de las letras se mantiene en el segundo y tercer nivel de la clasificación. Se ha construido un árbol filogenético que muestra la identidad en la secuencia de aminoácidos entre las proteínas codificadas por los genes *cry* y *cyt* conocidos. El primer nivel de la clasificación está limitado por un nivel de identidad aminoacídica del 45%, todas aquellas que presentan una identidad menor al 45% en su secuencia difieren en el primer rango (Cry1, Cry2, Cry3, etc.). Esto significa que proteínas con el mismo primer numeral, comparten al menos un 45% de similitud en su secuencia. El rango secundario indica una homología de al menos el 75% (Cry1A, Cry1B, Cry1C). Finalmente, las proteínas que comparten el primer numeral, letra mayúscula y letra minúscula comparten al menos un 95% de su secuencia de aminoácidos (Cry1Aa1, Cry1Aa2, Cry1Aa3, etc.) (De Maagd, Bravo y Crickmore 2001; Palma *et al.*, 2014).

Los genes *cry* tienen una secuencia codificante de entre 1900 pb y unas 3600 pb. Una característica que comparten, con muy pocas excepciones, es su expresión durante la fase estacionaria del cultivo. Las proteínas que codifican presentan una identidad aminoacídica que puede variar en gran medida desde más de un 90% hasta menos de un 20% y el intervalo de masas moleculares oscila entre 50 y 140 kDa (Evdokimov *et al.*, 2014).

Las proteínas Cry son propiamente protoxinas que se solubilizan en el medio alcalino del intestino de los insectos, y luego son proteolíticamente procesadas en un fragmento tóxico de menor tamaño. El extremo amino-terminal de la proteína (N_t) codifica para el dominio tóxico, mientras que el extremo carboxi-terminal (C_t) se cree que está relacionado con la formación y estabilización del cristal. Las dos partes de las proteínas también difieren en la conservación de las secuencias de aminoácidos. La parte C_t de las proteínas está altamente conservada, mientras que el extremo N_t presenta una mayor variación,

concentrándose ésta en regiones hipervariables que determinan la especificidad insecticida (Ge, Shivarova y Dean 1989). Sin embargo, esta región también presenta algunos bloques conservados.

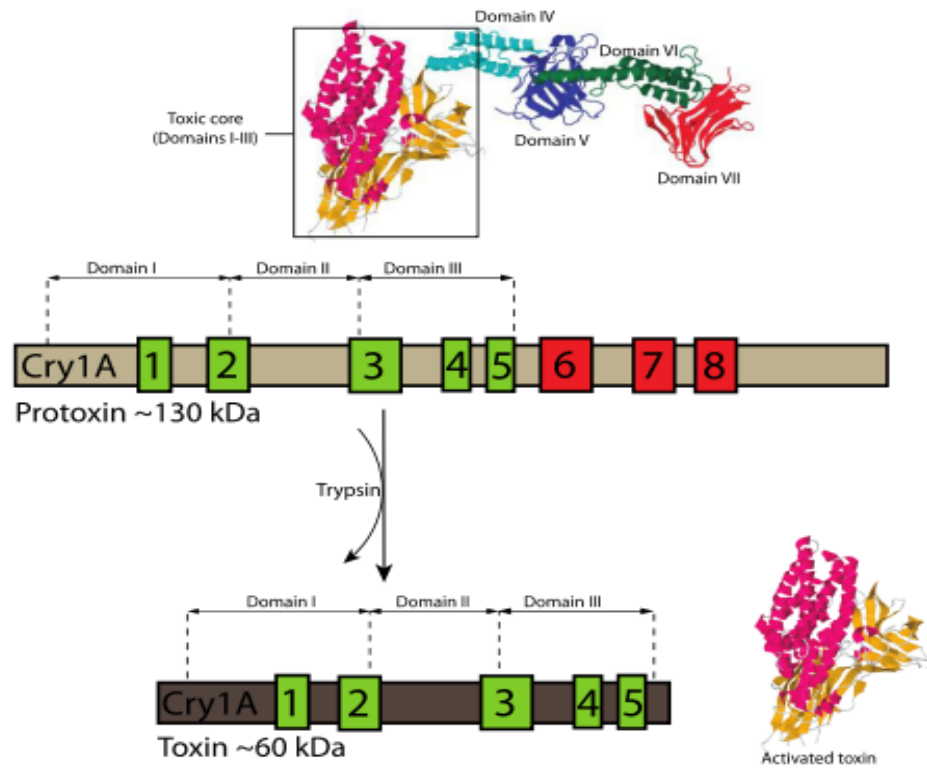


Figura 1. Bloques de aminoácidos conservados (1-8) entre las diferentes proteínas Cry de tres dominios (cajas de color). Cajas verdes representan los cinco bloques de aminoácidos conservados descritos por Höfte y Whiteley ubicados en el núcleo tóxico de la proteína Cry. Las cajas rojas indican los tres bloques de aminoácidos conservados adicionales encontrados por Schnepf *et al.* La protoxina es digerida por las proteasas del intestino medio en un fragmento más pequeño, que es responsable de la actividad tóxica. Cambios en la estructura tridimensional tras la activación proteolítica también se representan para la protoxina Cry1Ac: los dominios I, II y III comprenden el núcleo tóxico (en cajas verdes), el dominio IV (cian), el dominio V (azul), el dominio VI (verde), y el dominio VII (rojo).

Höfte and Whiteley (1989) en su revisión de la estructura de las proteínas Cry citaron la presencia de cinco bloques conservados en la región N_1 de las proteínas Cry entonces conocidas. Schnepf *et al.* (1998) identificaron estos bloques en la mayoría de las proteínas, junto con otros tres bloques adicionales en

el extremo C_i, fuera del núcleo toxico. Por otra parte, Crickmore (2000) ha señalado que estos bloques se encuentran en la mayoría de los genes *cry*. Un trabajo posterior ha revelado la estructura de la forma de la protoxina de Cry1Ac, demostrando que la región de la toxina se pliega en los tres dominios, mientras que la pro-región extendida forma otros 4 dominios. Los dominios IV y VI están formado por alfa hélices, mientras que los dominios V y VII tienen una topología formada por láminas β (Figura 1) (Evdokimov *et al.*, 2014).

Existen dos tipos de δ-endotoxinas: las proteínas Cry y las proteínas Cyt. Hasta la fecha se han clonado y secuenciado más de 800 genes *cry* diferentes. Las toxinas Cry que pertenecen a la familia de toxinas Cry de tres dominios, muestran claras diferencias en sus secuencias de aminoácidos, pero todos tienen en común una estructura de tres dominios muy similar y conservada (Figura 2) (Palma *et al.* 2014).

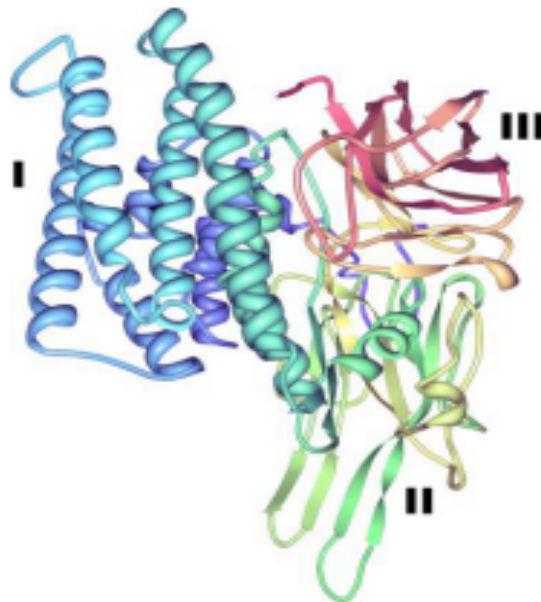


Figura 2. Estructura tridimensional de la toxina Cry2Aa. Esta estructura es representativa de una toxina de tres dominios producida por *Bt*. Números romanos indican los dominios típicos de las proteínas Cry de tres dominios: (I), dominio de perforación; (II), dominio central, involucrado participan en las interacciones de los receptores de la toxina; (III), dominio de unión a galactosa, participa en la unión al receptor y la formación de poros.

El dominio I o dominio de perforación, situada hacia el extremo N_t, está constituido por un grupo de siete α -hélices que se somete a la escisión proteolítica parcial en todas las proteínas Cry de tres dominios durante la activación de la toxina y esto puede ser responsable de la inserción en la membrana de la toxina y la formación de poros. El dominio II (dominio central o medio) consta de tres láminas de estructura β antiparalelas y juega un papel importante en las interacciones de los receptores de la toxina. Por último, el dominio III (dominio de unión a galactosa), es un β -sándwich en antiparalelo, éste está implicado en la unión al receptor y la formación de poros (Bravo, Gill y Soberon 2007; Sauka y Benintende 2008; Palma *et al.*, 2014).

2.4.4 Modo de acción de las proteínas Cry

Diferentes miembros de las toxinas Cry de tres dominios comparten plegamiento tridimensional similar, lo cual sugiere que también comparten un modo de acción similar. Las toxinas Cry de tres dominios son reconocidas por ser toxinas formadoras de poros que matan a las células del intestino medio del epitelio de las larvas provocando un choque osmótico y por consiguiente llevando a una lisis celular. El modo de acción de las proteínas Cry de tres dominios (Figura 3) ha sido estudiado principalmente en insectos lepidópteros. Hay tres diferentes modelos que han sido propuestos para explicar el modo de acción de las toxinas Cry de tres dominios: el modelo “clásico”, el modelo de “unión secuencial” y el modelo de “via de señalización” (Palma *et al.*, 2014).

El modelo clásico básicamente propone que la toxina lisa las células del epitelio del intestino medio de insectos susceptibles a través de los siguientes pasos: a) ingestión y disolución de la inclusión cristalina en el lumen del intestino medio alcalino; b) activación proteolítica de la protoxina que convierte la proteína Cry nativa en polipéptidos tóxicos más pequeños resistentes a las proteasas; c) unión de los fragmentos de la toxina activada a receptores específicos en la superficie de las células epiteliales del intestino medio; d) la formación de poros no

selectivos permeables a iones inorgánicos, aminoácidos y azúcares. Tales poros producen la lisis de las células epiteliales del intestino medio y por tanto desajustes que conducen a la muerte del insecto. Además, las esporas pueden colonizar, germinar, y replicarse en la hemolinfa, eventualmente matando las larvas por septicemia (De Maagd, Bravo y Crickmore 2001; Bravo, Gill y Soberon 2007; Sauka y Benintende 2008; Bravo *et al.*, 2013; Palma *et al.*, 2014; Melo, Soccol y Soccol 2016).

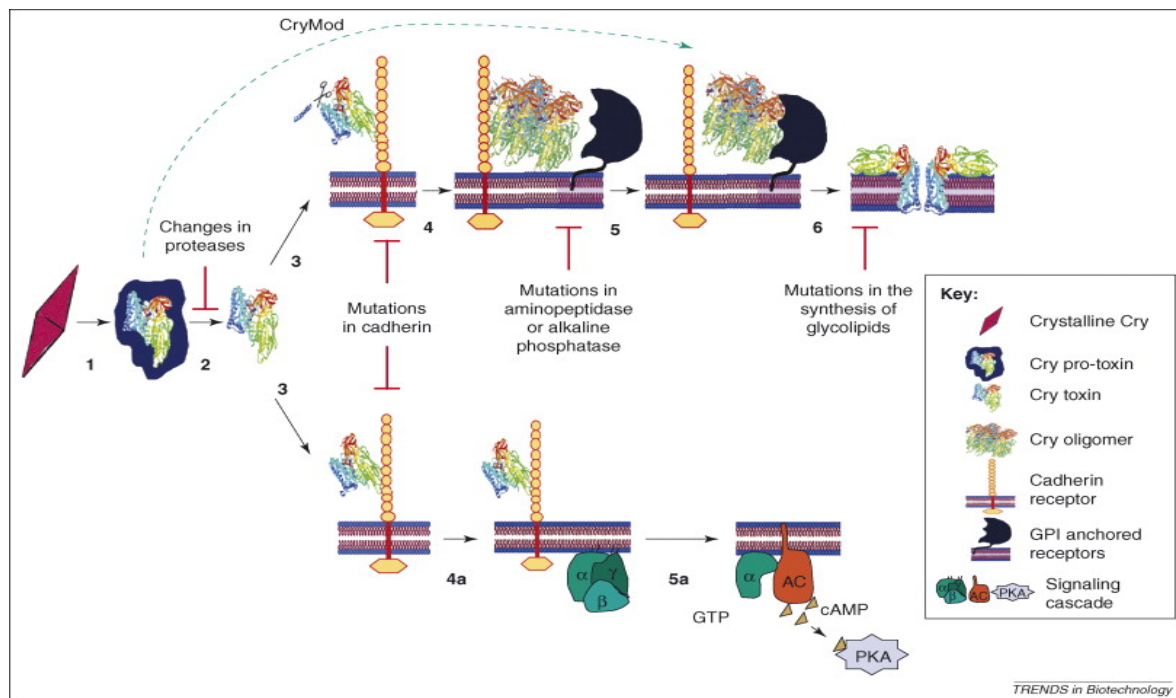


Figura 3. Modelos del modo de acción de las toxinas Cry y mecanismo resultante para la resistencia.

Se pueden distinguir dos mecanismos diferentes: el modelo de formación de poros (arriba) y el modelo de transducción de señales (abajo), que incluyen pasos iniciales similares para la solubilización de toxinas en la luz media del intestino (1), activación por proteasas del intestino medio (2) y unión al receptor primario de cadherina (3). En el modelo de formación de poros (arriba), el paso 3 induce la escisión de la hélice α -1 y desencadena la oligomerización de la toxina (4). El oligómero de la toxina se une a un receptor secundario, como la aminopeptidasa o la fosfatasa alcalina, que están ancladas por un ancla de glicosilfosfatidilinositol en la membrana (5). Finalmente, la toxina se inserta en la membrana, formando un poro que mata las células del insecto (6). El modelo de transducción de señales (abajo) propone que la interacción de la toxina Cry con un receptor de cadherina desencadena una vía de cascada intracelular que está mediada por la activación de la proteína G (4a), que en una etapa posterior (5a) activa la adenil ciclasa. Esta señal luego aumenta los niveles de monofosfato de adenosina cíclico, que activa la proteína quinasa A y conduce a la muerte celular. Las toxinas CryMod, en las que se elimina la hélice α -1, evitan la resistencia evitando la interacción de la cadherina.

El modelo de unión secuencial sugiere que las toxinas Cry, una vez activados por proteasas intestinales, se unen a proteínas tipo cadherina (glicoproteínas transmembranales que funcionan como receptores de toxinas) y se someten a un cambio conformacional que favorece la eliminación proteolítica de la hélice α -1 del dominio I y forma una estructura oligomérica llamada pre-poro. Más tarde, la unión a un receptor secundario, tal como una aminopeptidasa, facilita la inserción del pre-poro en la membrana, dando lugar a la muerte celular y por consiguiente del insecto (Palma *et al.*, 2014; Melo, Soccol y Soccol 2016).

En contraste, el modelo de la vía de señalización sugiere que la actividad tóxica está mediada por la unión específica a receptores de cadherina, dando lugar a una dependencia no descrita de canales de Mg^{2+} en la membrana plasmática, estos canales causan un movimiento anormal de iones en el citosol. El efecto de esto es mucho más grande debido a la destrucción de la membrana plasmática, y las toxinas bacterianas activan los procesos celulares pre-existentes durante la apoptosis como la adenilato ciclasa/proteína quinasa A, vía de señalización que produce la muerte celular necrótica (Palma *et al.*, 2014; Melo, Soccol y Soccol 2016).

2.4.5 Ecología de *Bacillus thuringiensis*

La ecología y el papel de *Bt* en la naturaleza han sido el objetivo de muchas preguntas aún sin respuesta. Algunos defienden que *Bt* es un microorganismo que vive en el suelo y que obtiene nutrientes para su supervivencia y reproducción en la naturaleza a partir de materia orgánica en descomposición o exudados de raíces, llegando a las partes aéreas de las plantas cuando éstas germinan y emergen del suelo. Una opinión opuesta afirma que *Bt* es un patógeno especializado que, al colonizar y matar a sus huéspedes y multiplicarse en sus cadáveres se deposita en el suelo y las plantas, que se convierten así en reservas naturales. Muchos autores clasifican a *Bt* como un microorganismo del suelo debido al hecho de que gran parte de su aislamiento a lo largo de los años se ha

obtenido a partir de este entorno (Fang *et al.*, 2011). Otros, sin embargo, consideran al suelo sólo como un medio de almacenamiento para las esporas de *Bt*, ya que apenas germinan en estos lugares, requiriendo nutrientes específicos y condiciones de pH específicas para eso. La microbiota nativa, así como las propiedades del suelo, como pH, humedad, minerales y concentraciones de materia orgánica tienen un efecto directo en la supervivencia de *Bt*, actuando positiva o negativamente sobre la germinación, crecimiento, esporulación y producción de proteínas. Estas condiciones, por otro lado, tienden a tener un efecto mucho menor sobre *Bacillus cereus* (un taxón muy relacionado con *Bt*), que ha demostrado la capacidad de multiplicarse en suelos no esterilizados, aumentando su población hasta un 20%. Sin embargo, bajo condiciones específicas, las esporas de *Bt* pueden germinar y crecer bien, como en suelos húmedos y ricos en nutrientes, con un pH cercano a la neutralidad, incluso en presencia de otras poblaciones microbianas. Además, las toxinas Cry de *Bt* pueden encontrar protección contra la degradación precoz en el suelo, ya que se asocian con ácidos húmicos y/o partículas de arcilla, lo que puede ayudar a mantener su toxicidad. (Argôlo-Filho 2014).

B. thuringiensis ha sido encontrado como un habitante natural del suelo de varias regiones del mundo. Las áreas sin historial de aplicación de productos basados en *Bt* han demostrado una gran diversidad de aislados de serotipo H con niveles variables de toxicidad, y la persistencia de *Bt* durante muchos años después de su aplicación en áreas de conservación ambiental también se ha informado. Petras y Casida (1985) notaron una reducción en el número de esporas en suelo, tanto en condiciones naturales como de laboratorio a través del tiempo, lo que concuerda con la idea de que *Bt* apenas se multiplica en este entorno, sin embargo, ha tendido a la estabilización. Por otro lado, la presencia de insectos no garantiza la presencia de *Bt* en el suelo. *Bt* se ha encontrado en suelos con poca o ninguna actividad de insectos, mientras que no se encontró en suelos con alta actividad de insectos (Martin y Travers, 1989). Además, *Bt* aparentemente no tiene

que ser un microbio obligado en su asociación con insectos, ya que crece en medios regulares *in vitro*.

Todavía no está claro cuál es la principal forma de dispersión de *Bt* en el suelo. Se sabe que su dispersión del suelo es limitada, ya que parece que muy poco de las poblaciones de *Bt* en el suelo se ve alterado por la acción de la lluvia, los invertebrados o el crecimiento de las plantas. Sin embargo, se ha descubierto que las esporas de *Bt* pueden pasar a través del tracto digestivo de vertebrados y organismos invertebrados, llegando a sus heces. En estas circunstancias, encuentran condiciones favorables para la germinación y, por lo tanto, llegan a ser capaces de dispersarse en el medio ambiente a través de la migración de estos animales (Zhang *et al.*, 2012).

Otro nicho ecológico del que se ha aislado *Bt* lo constituyen los insectos enfermos o muertos. La toxicidad de los aislamientos obtenidos a partir de ellos no se ajusta en todos los casos a lo esperable. Se han dado casos, por ejemplo, de aislamientos de *Bt* inactivos frente a la especie de la que fueron obtenidos o que afectan a insectos de distinto orden, lo que sugiere un posible comportamiento oportunista de la bacteria (Kaelin, 1994).

El filoplano constituye otro hábitat en el que es frecuente la presencia de cepas *Bt* con actividad insecticida. Los resultados obtenidos han llevado a postular que *Bt* podría ser una bacteria epífita de plantas, y que mantendría relaciones de simbiosis con ellas, aprovechando sus exudados para desarrollarse mientras ejerce un efecto repelente hacia los insectos. Además, la proporción de aislados con actividad insecticida contra lepidópteros provenientes de este medio es mucho mayor a la encontrada para aislamientos de suelo, lo que viene a apoyar la hipótesis de que la cubierta vegetal favorece el establecimiento de niveles enzoóticos de poblaciones de *Bt* activas frente a los insectos que la habitan. Existen otras muchas posibles fuentes de aislamiento, algunas de las cuales son de especial interés por su riqueza en *Bt*, como por ejemplo el polvo de almacenes

o molinos de cereal donde es particularmente frecuente y donde, curiosamente, abundan los aislados con actividad contra lepidópteros. Otro ejemplo, lo constituyen el agua estancada y los fondos de lugares anegados donde se desarrollan poblaciones de mosquitos. En ellos es frecuente el hallazgo de cepas de *B. thuringiensis* con actividad contra dípteros, explicable por la posible acumulación de residuos que contengan esporas generadas en insectos infectados (Irriarte y Caballero, 2001).

2.5 Identificación de genes *cry* en bacterias diferentes a *B. thuringiensis*

Existen varios reportes donde se han encontrado genes *cry* en otras especies de bacterias diferentes a *Bt* como *L. sphaericus*, *P. popilliae*, *C. bifermentans* subsp. *Malaysia* y *P. lentimorbus*.

En 1996 Barloy y cols, clonaron y expresaron el primer gen tóxico de *C. bifermentans* que codificaba una nueva proteína mosquitocida con homología a las delta-endotoxinas de *B. thuringiensis*. El gen *cry16Aa1* encontrado en *C. bifermentans* subsp. *malaysia* codifica una proteína mosquitocida de 71.128 kDa, esta toxina tipo Cry es la primera en encontrarse en un organismo anaerobio y la secuencia de aminoácidos deducida del gen es similar a las secuencias de las toxinas mosquitocidas de *B. thuringiensis*, especialmente en los primeros cuatro bloques conservados, por lo cual es probable que las toxinas mosquitocidas de *C. bifermentans* subsp. *malaysia* y *B. thuringiensis* tengan un ancestro en común.

En 1997 Zhang y cols. detectaron un cristal parasporal en extractos proteicos de esporangios de *P. popilliae* aislada de una larva enferma de *Melolontha melolonta* L. Un análisis de aminoácidos reveló homología significativa con la δ -endotoxina Cry2Aa de *B. thuringiensis*. El gen *cryBP1* (*Cry18Aa1*), que codifica para la proteína parasporal, fue encontrado en un operón *cry* en el cromosoma bacteriano, que contiene al menos otro marco de lectura abierto más pequeño, *orf1*. Los 706 aminoácidos de la proteína Cry18Aa1 tiene una masa

molecular predicha de 79 kDa y muestra aproximadamente el 40% de identidad de secuencia con las proteínas Cry2 de *B. thuringiensis*. Además del gen *cry18Aa1*, lograron identificar el gen *cry18Ba1* y *cry18Ca1* en *P. popilliae*

Barloy y cols. en 1998 reportaron la clonación y expresión de un segundo gen tóxico (*cry17Aa1*) que codifica una proteína de aproximadamente 71 kDa, también tiene algunas similitudes con los bloques conservados de la mayoría de las δ -endotoxinas de *B. thuringiensis*, pero no es significativamente tóxica para las larvas de mosquitos. Este gen (*cry17Aa1*) al igual que el gen *cry16Aa1*, también, fue identificado en *C. bifermentans* subsp. *malaysia* cepa CH18 que fue aislada del suelo manglar en Malasia por W.C. Cheong y H.L. Lee; por lo tanto, esta cepa es altamente toxica para larvas de mosquitos y su actividad se ve atribuida al gen *cry16Aa1*.

P. lentimorbus cepa Semadara, fue aislada de las larvas del escarabajo *Blitopertha orientalis* en la prefectura de Chiba, Japón, y mostró un alto nivel de actividad insecticida contra algunas especies de larvas de escarabajo. Esta cepa tiene la característica de formar esporangios en forma de huella de zapato con una gran espora terminal y un cuerpo parasporal más pequeño, esta cepa inhibe la ingestión y produce una secreción lechosa típica en el cuerpo de las larvas de escarabajo. La inhibición de la ingestión en *Anomala cuprea* se observa inmediatamente después de la inoculación, y se correlaciona con la concentración de esporangios, lo que sugiere que este fenómeno es debido al cristal proteico (Yokoyama *et al.*, 2004). Yokoyama y cols. (2004), seleccionaron una clona positiva de una biblioteca de DNA celular total de *P. lentimorbus* cepa Semadara que reaccionó con un antisuero que se generó contra las proteínas producidas por esta cepa. La clona positiva tenía un inserto de DNA que contiene dos genes *cry* (*cry43Aa1*, *cry43Ba1*), un gen *cry* parcial (*cry43-like*), y tres genes más pequeños situados río arriba. En las secuencias de aminoácidos de esta cepa se detectaron los ocho bloques que se conservan en las proteínas Cry de *B. thuringiensis*. Por otra parte, la cepa recombinante de *Escherichia coli* expresando el gen *cry43Aa1*

causó la inhibición de la ingestión y la mortalidad del 90% de larvas de *A. cuprea* de primer estadio larval.

Algunas cepas de *L. sphaericus* son altamente patógenas y producen proteínas Binarias (Bin) mosquitocidas; sin embargo, puede generarse resistencia a estas toxinas bajo condiciones de laboratorio. Jones y cols., en el 2007 analizaron cepas de *L. sphaericus* que pudieran superar esta resistencia revelando la presencia de un tipo de toxina de dos componentes que no habían sido descrita. Una subunidad Cry48Aa1 que es relacionada con las toxinas Cry de tres dominios de *B. thuringiensis*, únicamente para este tipo de proteína. La actividad insecticida es sólo activada en presencia de una segunda proteína, en este caso Cry49Aa1. Ambos componentes Cry48Aa1 y Cry49Aa1 son necesarios para la toxicidad, por lo tanto, indica una interacción sin precedentes para generar toxicidad. Esta nueva combinación de toxinas puede indicar una combinación fortuita de los miembros de las familias de genes que codifican toxinas Cry de tres dominios y las toxinas binarias.

2.6 Sistema de iniciadores universales.

Una de las técnicas más utilizadas para la amplificación de genes es el PCR. Esta técnica se ha utilizado en diferentes estudios para la identificación de genes *cry* proporcionando buenos resultados de manera rápida y precisa. Varios grupos de investigación han reportado el uso de PCR en la identificación de genes *cry*; sin embargo, hay deficiencias en su uso y se requiere utilizar muchos pasos o reacciones de PCR para el reconocimiento del gen, al igual que el empleo de varios iniciadores. Debido a esto, no es posible identificar genes nuevos que no estén relacionados a los ya conocidos.

En el 2010, Noguera e Ibarra diseñaron iniciadores universales a partir de alineamientos de los bloques conservados provenientes de los representantes de todas las familias de genes *cry* con el objetivo de identificar nuevos genes *cry*. En

este trabajo lograron diseñar cuatro iniciadores degenerados correspondientes a los bloques conservados 1, 2, 4 y 5 de las proteínas Cry de *Bt*. Realizaron un análisis *in silico* donde se estableció como estrategia el uso de tres grupos de amplificación o pares de iniciadores, con los cuales teóricamente se puede amplificar al 100% de los genes *cry* conocidos. Estos sets de iniciadores se probaron en 27 cepas atípicas de *Bt*, y los resultados revelaron que las amplificaciones son específicas y que los productos de amplificación en todos los casos correspondieron a genes *cry*. Después del análisis con Blast y de la clasificación de los genes, se lograron identificar 3 genes *cry* nuevos de primer rango a los que se les denominó *cry57Aa1*, *cry58Aa1* y *cry59Aa1* y 4 genes *cry* nuevos de segundo rango denominados como *cry8Ka2*, *cry8-like*, *Cry20Ba1* y *cry1Ma1* (Noguera e Ibarra 2010).

2.7 CD-HIT y su Algoritmo de agrupamiento

CD-HIT es un programa ampliamente utilizado para agrupar secuencias biológicas para disminuir la redundancia de secuencias y mejorar el rendimiento de otros análisis de secuencias. CD-HIT fue desarrollado originalmente para agrupar secuencias de proteínas para crear bases de datos de referencia con redundancia reducida (Li *et al.*, 2001) y luego se amplió para apoyar la agrupación de secuencias de nucleótidos y comparar dos conjuntos de datos (Li y Godzik 2006). El servidor web CD-HIT se implementó en 2009, lo que permite a los usuarios agrupar o comparar secuencias sin utilizar CD-HIT con una interfaz de línea de comandos o línea de texto simple. El servidor proporciona una interfaz interactiva y herramientas de visualización adicionales. Agrupar una base de datos de secuencias generalmente requiere comparaciones completas; por lo tanto, es muy lento. Muchos métodos usan BLAST para calcular las similitudes de todos contra todos. Es muy difícil para estos métodos agrupar grandes bases de datos. Mientras que CD-HIT puede evitar muchas alineaciones de secuencia con una palabra corta basada en heurística.

CD-HIT utiliza un proceso de agrupación incremental y evita las alineaciones innecesarias mediante un mecanismo de filtrado de palabras cortas. El objetivo de los filtros es decidir si la identidad entre dos secuencias está por encima o por debajo de un umbral sin alinearlas, lo que acelera el proceso de agrupamiento. Aunque CD-HIT se basa en la alineación, puede evitar demasiadas alineaciones por pares mediante el uso de un filtro, por lo tanto, es más rápido que BlastClust y puede manejar conjuntos de datos más grandes (Wei *et al.*, 2012).

El algoritmo básico CD-HIT clasifica las secuencias de entrada de largas a cortas y las procesa secuencialmente de la más larga a la más corta. La primera secuencia se clasifica automáticamente como la primera secuencia representativa del clúster. Luego, cada secuencia de consulta de las secuencias restantes se compara con las secuencias representativas que se encuentran antes de ella y se clasifica como redundante o representativa en función de si es similar a una de las secuencias representativas existentes. De manera predeterminada (modo rápido), una consulta se agrupa en el primer representante sin comparar con otros representantes. En el modo preciso, una consulta se compara con todos los representantes y se agrupa a la más similar.

Desde su lanzamiento, CD-HIT se ha utilizado ampliamente para una gran variedad de aplicaciones que van desde la creación de conjuntos de datos no redundantes (Suzek *et al.*, 2007), clasificaciones de familias de proteínas (Yooseph *et al.*, 2008), identificación de artefactos (Niu *et al.*, 2010), anotación metagenómica (Sun *et al.*, 2011), análisis de ARN (Loong y Mishra, 2007), a varios estudios de predicción (Rubinstein y Fiser, 2008).

2.8 HMMER

HMMER es un software que se utiliza para buscar bases de datos de secuencias para homólogos de proteínas o secuencias de DNA y para hacer alineamientos de secuencias. HMMER se puede utilizar para buscar en bases de

datos de secuencias con una secuencia de consulta única, pero se vuelve particularmente poderoso cuando la consulta es una alineación de múltiples instancias de una familia de secuencias. HMMER crea un perfil de la consulta que asigna un sistema de puntuación de posición específica para sustituciones, inserciones y eliminaciones. Los perfiles de HMMER son modelos probabilísticos llamados "Perfil de Modelos Ocultos de Markov" (HMM de perfil por sus siglas en inglés Hidden Markov Model) (Krogh *et al.*, 1994; Eddy, 1998; Durbin *et al.*, 1998).

Comparado con BLAST, FASTA y otras herramientas de alineación de secuencias y búsqueda de bases de datos basadas en metodología de puntuación más antigua, HMMER pretende ser significativamente más preciso y capaz de detectar homólogos, debido a la fortaleza de sus modelos de probabilidad subyacentes. En el pasado, esta fortaleza tenía un costo computacional significativo, con el perfil HMM las implementaciones corrían aproximadamente 100 veces más lentas que las búsquedas BLAST comparables para la búsqueda de proteínas, y aproximadamente 1000 veces más lentas que las búsquedas BLAST para la búsqueda de DNA. Con HMMER3.1, HMMER ahora es esencialmente tan rápido como BLAST para la búsqueda de proteínas, y aproximadamente 5-10 veces más lento que BLAST en la búsqueda de DNA.

Los perfiles HMM son modelos estadísticos de alineaciones de secuencias múltiples o incluso de secuencias únicas. Capturan información de posición específica sobre cómo se conserva cada columna de la alineación y qué residuos son posibles

2.8.1 Aplicaciones de perfiles HMM

HMMER se puede utilizar para reemplazar BLASTP y PSI-BLAST para buscar bases de datos de proteínas con secuencias de consulta únicas, a través de dos programas para este propósito: phmmer y jackhmmer, donde jackhmmer es una búsqueda iterativa similar a PSI-BLAST. Otra aplicación de HMMER

incluye trabajar con una familia de secuencias, con la cual se construye una alineación de secuencias múltiples. Esta familia de secuencias, como la mayoría de las familias de proteínas (o DNA), tiene una cantidad de aminoácidos clave (o nucleótidos) fuertemente conservados (pero no del todo), separados por espaciamientos característicos. Una de las complicaciones, es saber si hay más miembros de la familia de proteínas o DNA en las bases de datos de secuencias; sin embargo, debido a que la familia es evolutivamente diversa, una búsqueda con BLAST con cualquier secuencia individual no encontrará el resto de las secuencias que ya se conocen.

Las búsquedas de homología se usan ampliamente dentro de la biología molecular, facilitando la transferencia de la anotación desde una secuencia o región caracterizada funcionalmente a una región correspondiente en otra secuencia. Al buscar contra bases de datos de secuencias, el software HMMER utiliza perfiles de modelos ocultos de Markov (HMM) para representar la consulta, que puede tomar la forma de una sola secuencia de proteína o una secuencia de alineación múltiple. En el caso de una alineación de secuencia múltiple, las frecuencias de aminoácidos observadas en cada columna se convierten en probabilidades específicas de posición, con probabilidades por posición para ambas inserciones y eliminaciones, determinadas a partir de la alineación de entrada (Krogh *et al.*, 1994; Eddy 1998).

Los perfiles de modelos ocultos de Markov (perfil deHMM) (Durbin *et al.*, 1998; Krogh *et al.*, 1994) representan un avance importante en términos de sensibilidad para la búsqueda de secuencias con homología. Proporcionan un marco probabilístico formal para la comparación de secuencias y mejoran la detección de homólogos. En el pasado, esta sensibilidad mejorada tenía un costo computacional significativo, pero los avances recientes en HMMER3 han aumentado la velocidad de búsqueda de proteínas en al menos 100 veces, alcanzando velocidad tipo blastp a través de una combinación de filtros heurísticos (Eddy, 2008) e ingeniería informática (Eddy, 2011; Farrar, 2007).

2.8.2 Trabajos relevantes con el uso de perfiles HMM

El grupo de Bagos (2014) desarrolló un método basado en un Modelo Oculto de Markov, capaz de predecir las cadenas β transmembranales de las proteínas de la membrana externa de bacterias gram-negativas y discriminar aquellas de las proteínas solubles en agua en grandes conjuntos de datos. El modelo es enfocado de forma discriminatoria, con el objetivo de maximizar la probabilidad de predicciones correctas en lugar de la probabilidad de las secuencias. Además, el modelo HMM fue desarrollado únicamente con la secuencia de aminoácidos, sin el requisito de información evolutiva. El modelo es cíclico y captura las características estructurales de las cadenas β -transmembranales de las proteínas de la membrana externa. Para probar y evaluar el modelo, compilaron un conjunto de datos no redundante de 14 proteínas de membrana externa con una estructura conocida a resolución atómica. Además, el modelo es capaz de discriminar la membrana externa de las proteínas solubles en agua en aplicaciones a gran escala, con una tasa de éxito del 88.8% y 89.2% para la correcta clasificación de la membrana externa y las proteínas solubles en agua, respectivamente, las tasas más altas obtenidas en la literatura (Bagos *et al.*, 2014).

Por otro lado, en el grupo de Walsh (2017) examinaron la densidad y la distribución de la proteína lantibiótica subclase I (LanB), en conjuntos de datos del microbioma humano oral y de heces usando un perfil de Modelos Ocultos de Markov. El modelo se validó identificando correctamente los genes *lanB* conocidos en los genomas de bacterias productoras de lantibióticos de manera más efectiva que otros métodos, a la vez que era lo suficientemente sensible para diferenciar entre diferentes subclases de proteínas lantibióticas. Este enfoque se comparó con dos métodos existentes para detectar conjuntos de datos genómicos y metagenómicos obtenidos del Proyecto del Microbioma Humano (HMP). De los métodos evaluados, el nuevo perfil HMM identificó el mayor número de proteínas LanB putativas en los datos de heces y metagenoma oral, mientras que BlastP

identificó el menor número. Además, el modelo identificó más proteínas LanB que un modelo preexistente de Pfam lantionina deshidratasa. La búsqueda en el subconjunto del tracto gastrointestinal de la base de datos del genoma de referencia de HMP con el nuevo HMM identificó siete productores putativos de lantibióticos de la subclase I, incluidos dos miembros del género *Coprobacillus*. Estos hallazgos establecen que los perfiles HMM son una herramienta poderosa en la búsqueda de nuevos productores bioactivos.

Los perfiles HMM han sido utilizados para la búsqueda o identificación de proteínas, sin embargo, Gong y cols., en el 2012 utilizaron un perfil HMM para la caracterización de subtipos del gen de la hemaglutinina (HA) de la influenza A. El virus de la influenza A ha evolucionado a 16 subtipos de hemaglutinina (HA) con diferentes propiedades antigénicas. Hasta ahora, la tipificación se ha basado principalmente en análisis, pero las numerosas secuencias disponibles del Centro Nacional de Información Biotecnológica (NCBI) de EEUU ofrecen formas alternativas de caracterizar el gen HA. En este grupo de investigación lograron analizar todas las secuencias HA disponibles del NCBI. El paquete de software HMMER se usó para anotar cómo una secuencia de búsqueda encajaba en un perfil HMM construido a partir de la secuencia consenso de un subtipo HA particular.

Sharp *et al.*, (2017) desarrollaron un modelo bioinformático robusto que explota los modelos ocultos de Markov para la identificación de nucleasas bacteriocinas (NBs) en bacterias de las cuales, hasta la fecha, sólo se conocen unas pocas. Los NBs son toxinas grandes (> 60 kDa) que se dirigen a los ácidos nucleicos (DNA, ARNt o ARNr) en el citoplasma de bacterias susceptibles, por lo general estrechamente relacionadas con el organismo productor. Este grupo de investigación logró identificar más de 3000 genes NB localizados en plásmidos o en el cromosoma de 53 especies bacterianas distribuidas en diferentes nichos ecológicos, incluidos humanos, animales, plantas y el medio ambiente. Una nueva NB identificada se prevé que sería específica para *Pseudomonas aeruginosa*

(pyocin Sn) y se demostró que mata a *P. aeruginosa*, lo que validó el modelo desarrollado en su trabajo. Curiosamente, mientras que los genes que codifican la maquinaria necesaria para la translocación de NB a través de la envoltura celular están muy extendidos en las bacterias Gram-negativas, los NB se encuentran exclusivamente en las γ -proteobacterias. El análisis de redes de similitud demostró que los NB se dividen en ocho grupos, cada uno con una disposición distinta de los dominios de proteínas implicados en la importación. La única característica estructural conservada en todos los grupos fue un motivo de secuencia crítico para la destrucción celular que generalmente no se encuentra en las bacteriocinas que apuntan al periplasma, lo que implica un papel específico en la translocación de la nucleasa al citoplasma. Finalmente, demostraron una asociación significativa entre las nucleasas colicinas, los NB específicos para *E. coli* y los factores de virulencia, lo que sugiere que los NB desempeñan un papel en los procesos de infección, muy probablemente al permitir que los patógenos superen a las bacterias comensales.

Otro estudio en el cual hacen uso de los Modelos Ocultos de Markov es el desarrollado por Restrepo-Montoya y cols. (2011). En este estudio se describe un enfoque para la identificación a gran escala y selección de proteínas que desempeñan un papel activo en la invasión de reticulocitos humanos por *Plasmodium vivax*. Este estudio se basa en la identificación de proteínas secretadas y/o relacionadas con la invasión usando la reducción de la redundancia de las secuencias y perfiles HMM para analizar genes anotados en el genoma con niveles significativamente altos de transcripción hacia el final del ciclo intra-eritrocitario. La preferencia por esta etapa de la transcripción se basó en la premisa de que la expresión génica es oportunamente regulada, por lo tanto, las funciones biológicas pueden asignarse a un instante particular del ciclo de desarrollo del parásito dentro del glóbulo rojo infectado. Los candidatos seleccionados fueron analizados; esto incluyó revisar la evidencia experimental publicada de la función durante la invasión y/o ensayos inmunológicos que conducen a su clasificación como candidatos a vacunas y confirmación

experimental de presencia de proteína. El objetivo era definir criterios de selección sólidos basados sobre una estrategia bioinformática para encontrar proteínas de *P. vivax* que probablemente sean candidatas interesantes para la producción de vacunas.

3. JUSTIFICACIÓN

El control de insectos plaga y de vectores de enfermedades humanas se ha realizado principalmente con la aplicación de insecticidas químicos. Sin embargo, estos insecticidas han generado problemas de contaminación ambiental, toxicidad a insectos no-blanco y de manera importante a los agricultores que lo aplican. Por otra parte, los insecticidas químicos han perdido su eficacia en el control de insectos, ya que su aplicación ha generado la aparición de poblaciones de insectos resistentes. Debido a estos inconvenientes se ha utilizado por más de 50 años a *B. thuringiensis* para el control de insectos plaga de los órdenes Lepidóptera, Coleóptera, Díptera e Himenóptera) y nemátodos de forma segura para el ambiente y sin mostrar toxicidad contra mamíferos; sin embargo, tienen un rango muy estrecho de hospederos y no todas las cepas tienen alta toxicidad. Por ello se ha intensificado la búsqueda de nuevas proteínas Cry con el fin de identificar proteínas más tóxicas, diferentes o nuevas a las ya utilizadas dentro del control biológico, las cuales posean actividad insecticida contra insectos que no han podido ser controlados con *Bt* y constituyan una alternativa para el control de insectos que ya han desarrollado resistencia a dichas toxinas. Además, una búsqueda exhaustiva de estos genes en bacterias diferentes a *Bt* ayudará a dilucidar si los genes *cry* son exclusivos de *Bt* o si realmente se encuentran ampliamente distribuidos en diferentes especies de *Bacillus*.

4. HIPOTESIS

Existe una gran diversidad de genes *cry* conocidos en *B. thuringiensis* y hay evidencias sobre su presencia en otras bacterias, por lo tanto, es posible encontrar genes *cry* de tres dominios en otras especies de baciláceas diferentes a *B. thuringiensis*, mediante una búsqueda sistemática con el uso de técnicas moleculares y herramientas bioinformáticas, lo que permitiría elucidar la diversidad biológica de las proteínas Cry dentro de las baciláceas.

5. OBJETIVOS

5.1 General

Identificar genes *cry* de 3 dominios en bacterias diferentes a *B. thuringiensis* mediante el uso de iniciadores universales y mediante el uso de herramientas bioinformáticas, así como la caracterización y clasificación de los genes y de las cepas que los contengan.

5.2 Específicos:

- Seleccionar las cepas de *Bacillus* spp. y amplificar los posibles genes *cry* de tres dominios, clonar y secuenciar los amplicones obtenidos de las cepas seleccionadas, así como analizar y seleccionar los genes *cry* candidatos.
- Amplificar, clonar y secuenciar los genes *cry* completos obtenidos de *Bacillus* spp y caracterizar las cepas y los genes *cry* nuevos.
- Comprobar la expresión de los nuevos genes *cry* en cepas acristalíferas de *Bt* y realizar bioensayos cualitativos y/o cuantitativos de la cepa nativa y la recombinante con diferentes insectos y nemátodos.
- Construir una base de datos con las secuencias del C-terminal de las proteínas Cry de 3-dominios y realizar un perfil HMM, así como validar el perfil HMM con el plásmido pBtoxis de *B. thuringiensis israelensis* y con el genoma de *B. thuringiensis israelensis* cepa 4Q7.
- Construir una base de datos de los genomas reportados en el GenBank de género *Bacillus* y realizar una búsqueda de proteínas Cry con el perfil HMM y analizar los genomas positivos para genes *cry* con la técnica de tipificación de secuencia multilocus y determinar su especie.

6. MATERIALES Y MÉTODOS

6.1 Búsqueda de genes *cry* en cepas de *Bacillus* spp.

6.1.1 Selección de cepas y condiciones de crecimiento

En el presente trabajo se seleccionaron 223 cepas de bacilos bajo la única condición que no formaran cristal durante la fase de esporulación. Todas las cepas se observaron al microscopio de contraste de fases y una vez confirmada su carencia de cristal se procedió a la extracción del DNA genómico como se mencionará más adelante. Algunas cepas de bacilos fueron proporcionadas por el Laboratorio de Bacteriología Molecular de Cinvestav-Irapuato y el resto fueron obtenidas de la colección de *Bacillus* del Laboratorio de Bioinsecticidas de Cinvestav-Irapuato. Las cepas obtenidas del laboratorio de Bacteriología Molecular fueron cultivadas en caldo y agar Medio Marino (MM) [5% hidrolizado de caseína, 1% extracto de levadura, 85.55 mM NaCl, 7.04 mM Na₂SO₄, 3.60 mM CaCl, 2.68 mM KCl, 0.94 mM Na₂CO₃, 0.29 mM C₆H₅FeO₇, 40 mM MgSO₄, 0.67 mM KBr, 0.35 mM H₃BO₃, 0.57 mM NaF, 0.19 mM NO₃NH₄, y 0.56 mM Na₂HPO₄] (Gómez-Lunar *et al.*, 2016). El resto de las cepas fueron cultivadas en caldo LB y agar utilizando métodos estándar (Difco, Becton, Dickinson, Mexico) e incubadas a 28°C.

E. coli (DH5α) fue usada para la transformación estándar con los vectores pCR 4-Topo® (Invitrogen, Carlsbad, California), pJET1.2/blunt (Thermo Fisher) y pSTAB. Las transformantes de *E. coli* fueron cultivadas en caldo y agar Luria-Bertani (LB) [1 % triptona, 0.5 % extracto de levadura y 1 % NaCl; pH 7.0] por 24 horas a 37°C con una agitación de 200 rpm.

La cepa acristalífera de *B. thuringiensis* subsp. *israelensis* 4Q7 y la cepa acristalífera *B. thuringiensis* subsp. *kurstaki* HD1 Cry^B fueron utilizadas para la transformación del vector pSTAB conteniendo el gen *cry*. Las transformantes de *Bt* fueron cultivadas en caldo LB y agar utilizando métodos estándar (Difco, Becton,

Dickinson, Mexico) e incubadas a 28°C a 250 rpm por 24 hrs. Ampicilina (100 µg/mL) o eritromicina (50 µg/mL) fue agregado al medio cuando fue apropiado, para la selección de cepas de *E. coli* y *Bt* resistentes.

6.1.2 Extracción de DNA genómico

Se extrajo DNA de cada cepa, siguiendo un protocolo modificado (Reyes Ramírez e Ibarra 2005). Se inoculó cada cepa en medio líquido a 28°C con agitación a 250 rpm toda la noche. Se centrifugó 1.5 ml del medio a 13000 rpm (Heraeus™ Pico™ 21 Microcentrifuge) por 5 min y se desechó el sobrenadante. Se resuspendió la pastilla en 300 µL de regulador J (Tris-HCl 1M, EDTA 0.1M, NaCl 0.15 M [pH8]) y se centrifugó a 13000 rpm (Heraeus™ Pico™ 21 Microcentrifuge) por 5 min. Se desechó el sobrenadante y la pastilla se resuspendió en 200 µL de regulador J conteniendo 12 µL de lisozima (40 mg/mL), se incubó la muestra a 37°C por 30 min. Se adicionaron 12 µL de SDS al 10% y se incubó a 70°C por 20 min, transcurrido el tiempo se bajó la temperatura a 50°C. Posteriormente se agregaron 5 µL de RNAsa (10 mg/mL) y se incubó 50°C por 30 min. Se adicionaron 7 µL de proteinasa K (10 mg/mL) y se incubó a 60°C por 1.5 horas, después se agregaron 35 µL de NaCl 5 M, se incubó en hielo por 15 min y se centrifugó a 13000 rpm (Heraeus™ Pico™ 21 Microcentrifuge) por 10 min. Por último, se recuperó el sobrenadante en tubos estériles y se precipitó con 300 µL de isopropanol frío, se incubó a -20°C por 30 min y se centrifugó a 13000 rpm (Heraeus™ Pico™ 21 Microcentrifuge) por 10 min, se eliminó el sobrenadante y se lavó la pastilla con etanol al 70%, se secó la pastilla y se disolvió el DNA en 30 µL de Tris-EDTA (pH 8). El DNA fue cuantificado y almacenado a -20°C hasta su uso.

6.1.3 Amplificación de los posibles genes *cry*

Para la búsqueda de genes *cry* se utilizaron tres pares de iniciadores (Tabla 1) reportados por Noguera e Ibarra (2010), los cuales se diseñaron a partir de los bloques conservados de los genes *cry*. El primer par estuvo diseñado para

amplificar una secuencia entre el bloque 1 y 5, el segundo par se diseñó para amplificar la región entre el bloque 2 y 5 y el tercer par amplifica la región entre el bloque 1 y 4. Las amplificaciones fueron llevadas a cabo bajo las siguientes condiciones: 2.5 μ L de 10X PCR buffer, 1 μ L de la mezcla de 10 mM dNTPs, 0.75 μ L de 50 mM MgCl₂, 1 μ L de iniciador 10 μ M, 100 ng de DNA, 0.5U de platinum TaqDNA Polymerasa (Invitrogen) y agua a un volumen final de 25 μ L. La reacción de amplificación fue llevada en un Veriti Thermal Cycler (Applied Biosystems) bajo las siguientes condiciones: una desnaturalización inicial de 94°C por 2 min, seguido por 30 ciclos de desnaturalización a 94°C por 30 s, alineamiento a 45°C por 30 s, la polimerización a 72°C por 1.5 min, por último, una extensión a 72°C por 10 min. Las muestras fueron analizadas en geles de agarosa al 1%.

6.1.4 Clonación, transformación, secuenciación e identificación de los posibles genes *cry*.

Debido a que las cepas de *B. thuringiensis* usualmente contienen más de un gen *cry* y diferentes amplicones muestran tamaños similares, cada producto de amplicón fue clonado en el vector Topo PCR4 (Invitrogen) y transformado en *E. coli* (DH5 α) siguiendo las indicaciones del fabricante. Posteriormente se hizo un análisis de restricción de las clonas obtenidas usando las enzimas *EcoRI*, *EcoRV*, *HindIII*, *BamHI*, y *AluI* para identificar diferentes secuencias *cry* en las clonas transformadas. Los amplicones fueron secuenciados (Macrogen, Co.) y analizados usando el programa Blastn y Blastx del NCBI para identificar los genes parciales de las proteínas Cry.

6.1.5 Obtención e identificación de la secuencia completa del gen *cry*

Para la obtención de la secuencia completa del gen *cry* se utilizó la técnica de Genome Walker universal kit (Clontech). Para llevar a cabo esta técnica se siguieron las especificaciones del fabricante y se diseñaron 4 iniciadores (Tabla 1). Se diseñaron dos iniciadores reversos específicos del gen, estos se diseñaron a

partir del extremo 5' de la secuencia conocida, uno para la PCR primaria (GSP1Rv) y el otro para la PCR secundaria (GSP2Rv); los otros dos iniciadores directos (GSP1Fw y GSP2Fw) se diseñaron a partir del extremo 3' de la secuencia conocida del gen, con la cual se obtuvo la secuencia faltante del gen tanto río arriba como río abajo.

A partir de estas secuencias se diseñaron los iniciadores para la amplificación completa del gen *cry* (Tabla1). Además, a estos iniciadores se les agregó los sitios *Sall* y *SphI* para la clonación final en el vector de expresión pSTAB. La amplificación completa del gen fue llevada a cabo bajo las siguientes condiciones: 10 µL de 5X Phusion® HF Buffer, 1 µL de 10 mM dNTPs, 2 µL de cada iniciador 10 µM, 10 ng de DNA, 0.5µL de Phusion High-Fidelity DNA Polymerase (Invitrogen) y agua a un volumen final de 50 µL. La reacción de amplificación fue llevada en un Veriti Thermal Cycler (Applied Biosystems) bajo las siguientes condiciones: una desnaturalización inicial de 98°C por 30 s, seguido por 30 ciclos de desnaturalización a 98°C por 10 s, alineamiento a 55°C por 30 s, la polimerización a 72°C por 1 min, por último, una extensión a 72°C por 10 min. Las muestras fueron analizadas en geles de agarosa al 1%.

Cada producto del PCR fue clonado en el vector pJET1.2/blunt (Thermo Fisher) y transformado en la cepa DH5α. Se seleccionaron varias clonas y se les extrajo el DNA plasmídico, éste fue digerido con las enzimas *BglII*, *EcoRV*, *Sall* y *SphI*. Esta construcción fue llamada pJET-cry-like, el amplicón correspondiente al gen *cry-like* fue secuenciado (Macrogen, Co.) y analizado usando el programa Blastx del NCBI para su identificación.

Posteriormente, el amplicón del gen *cry-like* se clonó en el vector de expresión pSTAB utilizando los sitios *Sall* y *SphI*, por último, se transformó en DH5α. El DNA plasmídico obtenido de las clonas transformadas fue analizado con la enzima *EcoRI* y una doble digestión con *Sall* y *SphI* para corroborar la apropiada inserción del gen *cry-like* en pSTAB. Este nuevo plásmido fue llamado

pSTAB-*cry-like*. Se secuenció nuevamente utilizando los iniciadores de amplificación del gen, así como tres iniciadores internos de la secuencia, por último, ésta se depositó en la base de datos del GenBank. La identidad con otros genes *cry* y los análisis de alineación se llevaron a cabo utilizando el programa Blastn y Blastx del NCBI (Zhang *et al.*, 2000). Para la predicción 3D de la estructura de la δ -endotoxina se utilizó el modo intensivo del software Phyre2 como lo describe Kelley (2015) y Tabassum (2017).

6.1.6 Transformación de las cepas acristalíferas de *Bt*

Células competentes de la cepa 4Q7 y Cry^B fueron electroporadas con 3 μ L de DNA plasmídico usando los siguientes parámetros: 2000 V, 335 Ω y 25 μ F. Las células fueron cultivadas en 800 μ L de medio LB a 28°C por dos horas antes de que fueran cultivadas en cajas de agar LB conteniendo 50 μ g/mL de eritromicina e incubadas a 28°C. Se seleccionó una clona de cada cepa transformante y se corroboró la transformación por PCR y secuenciación.

6.1.7 Análisis de proteínas

Se realizó una electroforesis en gel de poliacrilamida SDS (SDS-PAGE) como lo describe Laemmli (1970) usando un gel al 10%. La electroforesis se realizó a 30 mA durante 0.5 horas y 100 mA durante 5 horas. Los geles se tiñeron con Coomassie Brilliant Blue R-250 y se destiñeron en ácido acético glacial al 10% (v/v) y metanol al 10% (v/v). PageRuler Protein Ladder (Promega, Fitchburg, Wisconsin, USA) se usó como marcador molecular. La expresión de la nueva proteína Cry-like tanto en la cepa LBIC-004 y 4Q7-cry-like se analizó mediante desorción láser asistida por matriz-tiempo de vuelo de espectrometría de masas (MALDI-TOF MS) según lo descrito por Noguera e Ibarra (2010) y Hernández-Soto y *col.* (2009).

6.1.8 Caracterización de la cepa

Para la identificación taxonómica se realizaron exámenes morfológicos preliminares mediante microscopía de contraste de fases. Posteriormente, diferentes genes fueron amplificados y secuenciados. Estos incluyen el gen *16S rRNA* que codifica para la subunidad pequeña del RNA ribosomal y el gen *gyrB* que codifica para la subunidad B de la topoisomerasa tipo I, además se usó el gen de la flagelina (*hag*) para complementar el análisis filogenético (Reinoso-Pozo *et al.*, 2016; Xu y Côté, 2006; Zhang *et al.*, 2000; Soufiane y Côté, 2009). Los iniciadores y las temperaturas de alineamiento utilizados para la amplificación de estos genes se muestran en la Tabla 1. La caracterización adicional de la cepa incluyó el patrón de plásmidos. Para el patrón plásmidos se siguió la metodología descrita por Reyes-Ramírez e Ibarra (2008) y Reinoso-Pozo y cols. (2016).

Tabla 1. Características generales de los iniciadores utilizados para la caracterización de la cepa LBIC-004 y el gen *cry-like*.

Primer	Secuencia	Gen reconocido	Tamaño del producto (pb)	Temperatura de alineamiento (°C)
Bloque 1	5'-TATGCWCAAGCWGCCAATYTWCATYT-3'	Genes <i>cry</i>	1300	45°C
Bloque 5	5'-GGRATAAATTC AATTYKRTCWA-3'	Genes <i>cry</i>	1300	45°C
GSP1Rv	5'-GTAATCAACCCAGTTCTTAGCAGATGATCC-3'	<i>cry-like</i>	Inespecifico	67°C
GSP2Rv	5'-GCCAGTATGATACCATTCTACACAATG-3'	<i>cry-like</i>	Inespecifico	67°C
GSP1Fw	5'-CGTATGAGGATTAGATATGCCGCTGATTC-3'	<i>cry-like</i>	Inespecifico	67°C
GSP2Fw	5'-CAGATCTACATATCAAATGCCCTTTAAAGCAAC-3'	<i>cry-like</i>	Inespecifico	67°C
Cry8Fw	5'-ATGAGTCCRAATAATCARAATGART-3'	<i>Cry8</i>	ND	50°C
Cry8Rv	5'-TTACTCTTCTTCTAACACGAGTTCTACAC-3'	<i>Cry8</i>	ND	50°C
Cry-likeFw	5'-GCGCGTCGACATGAATTC AAAAGACGAAAACAAAC-3'	<i>cry-like</i>	1953	55°C
Cry-likeRv	5'-CGCAGCATGCTTACATAGGTATAAACTCAATTTTATC-3'	<i>cry-like</i>	1953	55°C
fD1	5'-AGAGTTTGATCCTGGCTCAG-3'	<i>16s rRNA</i>	1400	55°C
rP2	5'-ACGGCTACCTTGTTACGACTT-3'	<i>16s rRNA</i>	1400	55°C
gyrB-F1	5'-ATGGAACAAAAGCAAATG-3'	<i>gyrB</i>	2200	46°C
gyrB-Fr1	5'-TTAAATATCAAGGTTTTTCA-3'	<i>gyrB</i>	2200	46°C
BThagF1	5'-AGTACATGCGCCAAAACCAAG-3'	<i>flagelina</i>	750	54°C
BThagR1	5'-GTTTGCTTGAGAAAGCATGCT-3'	<i>flagelina</i>	750	54°C

ND: no determinado

6.1.9 Bioensayos

Se realizaron bioensayos con la cepa LBIC-004, la cual contenía un gen *cry-like* y la cepa transformada con el gen *cry-like* (4Q7-*cry-like*) contra insectos y nemátodos. Los bioensayos de toxicidad hacia insectos incluyeron larvas de cuarto instar temprano de *Aedes aegypti*, larvas neonatas de 24 horas de *Manduca sexta* y larvas de *Phyllophaga* sp. Los ensayos de actividad de nemátodos se realizaron contra *Caenorhabditis elegans* en segundo estado juvenil (J2).

Los bioensayos cualitativos contra larvas de *A. aegypti* se realizaron en recipientes conteniendo 100 mL de agua desclorada con 20 larvas de *A. aegypti* agregando una asada de la cepa LBIC-004 y de la 4Q7-Cry-like. Como control positivo se utilizó a *B. thuringiensis* subsp. *israelensis*. Cada ensayo se hizo por triplicado, las larvas fueron incubadas a 28°C y examinadas después de 24 horas (Ibarra *et al.*, 2003). Para los bioensayos con *M. sexta* se utilizaron 20 larvas de primer instar y se colocaron en placas Petri de 60 cm² conteniendo dieta artificial y una concentración de 10 µg/cm² para cada tratamiento. Para este bioensayo se utilizaron seis tratamientos: la cepa en estado vegetativo, en estado de esporulación y células lisadas, así como los sobrenadantes de cada estado de crecimiento. Como control positivo se utilizó a *B. thuringiensis* var. *kurstaki* y cada ensayo se hizo por triplicado (Barboza-Corona *et al.*, 1998).

El bioensayo realizado con *C. elegans* (~30 individuos) se llevó a cabo por 7 días y se utilizaron 100 µL/mL de solución de esporangio. Para este bioensayo se utilizaron como control negativo a *E. coli* cepa OP50 y *B. thuringiensis* var. *kurstaki*. Como control positivo se utilizó la cepa LBIT-107, con capacidad nematocida, y se realizó por triplicado. Por último, se realizó el bioensayo contra *Phyllophaga* sp., para este bioensayo se utilizaron 100 µg de esporangio por larva en 3 µL de solución de Tween 80 al 0.02% y sacarosa al 20%. Se utilizaron 10 larvas por tratamiento y el bioensayo se llevó a cabo por 5 días.

6.2 Búsqueda *in silico* de proteínas Cry en genomas de *Bacillus* spp

6.2.1 Recuperación de secuencias de las proteínas Cry

Se obtuvieron todas las secuencias posibles de las proteínas Cry de tres dominios (http://www.lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/) y se descargó la secuencia de aminoácidos de cada proteína disponible en el National Center for Biotechnology Information (NCBI) hasta Enero del 2017.

6.2.2 Agrupación y comparación de secuencias con CD-HIT

Para la agrupación de las secuencias de las proteínas Cry se utilizó el programa de CD-HIT. Este ha sido extensamente utilizado para el agrupamiento y comparación de proteínas y secuencias de nucleótidos. Para agrupar las secuencias de aminoácidos de las proteínas Cry se utilizó el paquete de software cd-hit-v4.6.7 (<https://github.com/weizhongli/cdhit/releases>), los agrupamientos se realizaron en un inicio desde un 50% hasta un 90% de similitud en las secuencias proteicas y se observó su comportamiento analizando cada clúster de forma manual mediante el uso del programa MEGA7. El algoritmo básico CD-HIT clasifica las secuencias de largas a cortas, la primera secuencia se clasifica automáticamente como la primera secuencia representativa del clúster; después cada secuencia de consulta de las secuencias restantes se compara con las secuencias representativas que se encuentran antes de ella y se clasifica como redundante o representativa en función de si es similar a una de las secuencias representativas existentes. La ventaja de este programa es su velocidad ultra rápida y que puede evitar muchos alineamientos de secuencia con filtros de palabras cortas (Huang *et al.*, 2010; Fu *et al.*, 2012; Weizhong *et al.*, 2012). Los algoritmos detallados para CD-HIT se pueden encontrar en trabajos anteriores (Li, Jaroszewski y Godzik 2001, 2002; Li y Godzik 2006). Finalmente, antes de realizar el perfil HMM se realizó un agrupamiento con el paquete de software cd-hit-v4.6.7 con un 70% de similitud entre las secuencias de las proteínas Cry, para realizar el

agrupamiento se utilizó el script CD-HIT-PARA, el cual está escrito en el lenguaje de programación Perl, por último, los agrupamientos fueron analizados con MEGA7; además, se tomaron en cuenta otros criterios que se mencionan más adelante para hacer el modelo más estricto.

6.2.3 Construcción del perfil HMM

Para construir el perfil HMMs de las proteínas Cry se utilizó el paquete de software HMMER versión 3.1b2 (Eddy 1998). Las secuencias de las proteínas Cry fueron previamente agrupadas con el programa CD-HIT-V4.6.7 con un 70% de similitud y posteriormente el resultado de los múltiples agrupamientos y alineamientos de secuencias fue analizado por el programa MEGA7 como se mencionó anteriormente. Para el agrupamiento con un 70% de similitud se incluyeron varios criterios:

- Criterio de exclusión: Todas las secuencias de proteínas Cry que tuvieran su secuencia parcial o incompleta y aquellas proteínas que estuvieran clasificadas como δ -endotoxinas pero que no forman los tres dominios característicos de las proteínas Cry de tres dominios (Cry6Aa, Cry15Aa1, Cry22, Cry23, Cry33, Cry34, Cry35, Cry36, Cry37Aa1, Cry38, Cry45, Cry46 y Cry51Aa1), fueron excluidas del análisis.
- Criterio de eliminación: Todas las secuencias de proteínas Cry que son naturalmente truncadas, es decir que carecen en su forma natural del extremo carboxilo, también fueron eliminadas, ya que el modelo se construyó a partir de este extremo de las proteínas que se encuentra altamente conservado. En este punto se seleccionaron todas las secuencias de las proteínas Cry que tuvieran más de 800 aminoácidos en su secuencia a excepción de dos secuencias (GenBank: AAX53094.1 y CAA31951.1) que tenían más de 800 aminoácidos pero que eran más cortas que el resto de las secuencias e interferían con el alineamiento del extremo carboxilo terminal.

- Criterio de inclusión: De los clústeres obtenidos de los criterios anteriores sólo se seleccionaron los clústeres que tuvieran más de 5 secuencias alineadas.

El resultado del alineamiento con un 70% de identidad que incluye estos criterios fue leído por el algoritmo *hmmbuild* del programa HMMER para crear un perfil HMM. Este algoritmo nos generó nuestro archivo con la extensión *.hmm* (en este caso, *Cry.hmm*) que contiene la secuencia consenso para nuestra familia de proteínas Cry. Posteriormente se utilizó el programa *hmmsearch* para buscar proteínas de la familia Cry en la base de datos de *Bacillus* (previamente elaborada y descargada para nuestro trabajo) mediante el modelo desarrollado por HMM, para el cual se incluyeron dos criterios: mostrar solo aquellas proteínas que tengan un e-value menor a 1×10^{-5} y con más del 30% de similitud (Gong, Chen y Shih 2012; Muñoz-Medina *et al.*, 2015).

6.2.4 Validación del perfil HMM y análisis de los datos obtenidos por el modelo

Para la validación del modelo se utilizó como control positivo el plásmido pBtoxis de *B. thuringiensis israelensis*, el cual tiene 4 proteínas Cry y sólo dos de ellas tienen el extremo C-terminal (Cry4Aa y Cry4Ba) (Berry *et al.*, 2002) y como control negativo se utilizó *B. thuringiensis israelensis* cepa 4Q7 (Haeyoung *et al.*, 2014). Por otro lado, ya que se contaba con una base de datos de *Bacillus* obtenida del GenBank, la mayoría de las proteínas que el modelo HMM identificó se encontraban anotadas con su descripción correspondiente a proteínas Cry de tres dominios, a excepción de 49 anotaciones que tuvieron que ser revisadas manualmente en la base de datos del NCBI utilizando la herramienta Blastp para su clasificación, ya que estaban anotadas como proteínas hipotéticas o productos proteicos sin nombre.

6.2.5 Análisis MLST y reconstrucción filogenética

Con el objeto de confirmar o reconsiderar la identificación de las especies de *Bacillus* registradas en la base de datos del NCBI, se analizaron todos los genomas con secuencias Cry por medio del Multilocus Sequence Typing (MLST). Todos los perfiles alélicos o tipo de secuencia (Sequence type: ST) del análisis de MLST de los genomas usados en este trabajo que fueron obtenidos del GenBank se obtuvieron de PubMLST.org y algunos se enviaron al servidor MLST-1.8 del Center for Genomic Epidemiology (CGE). Las secuencias concatenadas de los genes del MLST se analizaron usando el programa MEGA7 (Tamura *et al.*, 2011) y la herramienta en línea iTOL: Interactive Tree of Life (Letunic y Bork, 2016). El análisis incluyó valores de arranque con 1000 réplicas.

7. RESULTADOS

En el presente estudio se seleccionaron 223 cepas con morfología similar a *B. thuringiensis*, pero ninguna mostró la capacidad de formar cristal durante la fase de esporulación. Entre las principales especies de bacilos con las que se trabajó fueron *B. subtilis*, *B. pumilus*, *B. mariflavi*, *B. cereus*, *B. horikoshii*, *B. megaterium*, *B. spp* y *L. sphaericus* (antes *B. sphaericus*). Una vez que las condiciones de amplificación fueron establecidas, el DNA de las 223 cepas fue sometido a la amplificación por PCR usando el sistema de tres sets de iniciadores reportado por Noguera e Ibarra (2010). Solo 32 cepas de las 223 analizadas mostraron al menos un producto de amplificación obtenido con cualquiera de los tres sets. De estas 32 cepas se lograron obtener 45 amplicones con los tres sets de iniciadores (Figura 4).

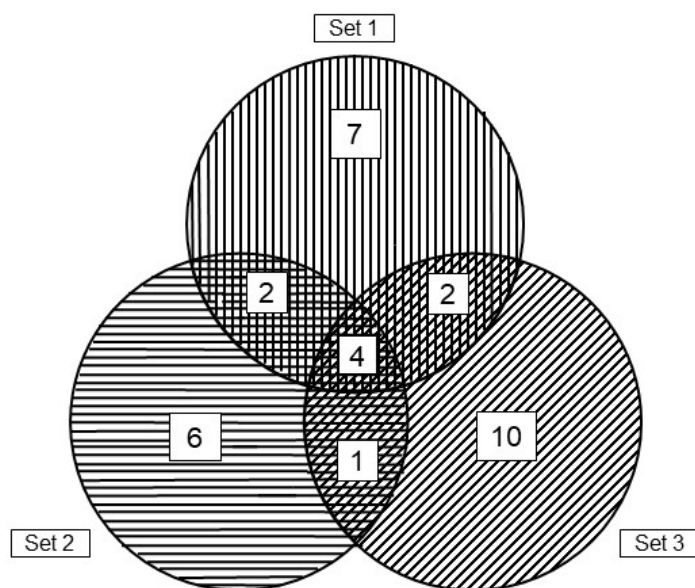


Figura 4. Diagrama de Venn de tres conjuntos de cepas que amplifican con cualquiera de los tres sets de iniciadores. El número de set indica la combinación de iniciadores utilizados en la amplificación como esta reportado por Noguera e Ibarra (2010).

Con el Set 1, 15 cepas mostraron un producto de amplificación de 1300 pb aproximadamente, con el Set 2, sólo 13 cepas mostraron productos de amplificación de 1000 a 1200 pb aproximadamente y con el Set 3, 17 cepas mostraron productos de amplificación de 1000 a 1200 pb. Como se muestra en el diagrama de Venn solo 4 cepas mostraron un producto de amplificación con cualquiera de los tres sets de iniciadores, 6 cepas mostraron un producto de amplificación con el Set 1 y Set2 y otras 6 cepas con el Set1 y Set3 y solo 5 cepas amplificó con el Set2 y Set3.

De los 45 amplicones obtenidos, 13 fueron seleccionados para ser clonados ya que mostraron la mayor probabilidad de ser genes *cry*. Esta selección se basó en el tamaño de amplicón esperado, dando prioridad a aquellos amplificados por el conjunto 1 que teóricamente es capaz de amplificar el 90% de los genes *cry* conocidos. Además, se incluyeron los amplicones obtenidos con la combinación de los conjuntos.

Los amplicones fueron clonados en el vector Topo pCR4 (Invitrogen, Carlsbad, California), cuyos productos de transformación fueron analizados con diferentes enzimas de restricción (*EcoRI*, *EcoRV*, *HindIII*, *BamHI*). En este caso, todos los amplicones mostraron el mismo patrón de restricción lo cual indicó que se trataba posiblemente de un solo gen por amplicón. Los amplicones clonados fueron secuenciados por la técnica de Sanger. Las secuencias fueron analizadas usando los programas del NCBI como BLASTn, NCBI BLASTp y NCBIx con las opciones predeterminadas. De estos 13 amplicones que fueron analizados sólo 6 mostraron un 65% de similitud con los genes *cry* (Tabla 2). Ya que las secuencias de las cepas LBIC-R6, LBIC-S9, LBIC-T4, LBIC-T7, LBIC-Y4 y LBIC-004 mostraron un 65% de similitud con la proteína Cry8, se realizó un alineamiento para corroborar que se trataba posiblemente del mismo gen, mostrando exactamente la misma secuencia (Figura 5), además, los resultados obtenidos del análisis del gen *16S rRNA* y del gen *gyrB* de las 6 cepas (Anexo 1) mostró que

estas cepas eran sólo réplicas del mismo aislado, por lo que se decidió utilizar sólo la cepa LBIC-004, para estudios posteriores.

```

LBICR6  CGAGGGAGATATATACGGATCCGGTGGGGAGTTATTTACCTACAAAGGTAATTTTATGATGTTATGTCCTGGTATGAGATGACAAAGGCATATCGCCAACTAC
LBICS9  CGAGGGAGATATATACGGATCCGGTGGGGAGTTATTTACCTACAAAGGTAATTTTATGATGTTATGTCCTGGTATGAGATGACAAAGGCATATCGCCAACTAC
LBICT4  CGAGGGAGATATATACGGATCCGGTGGGGAGTTATTTACCTACAAAGGTAATTTTATGATGTTATGTCCTGGTATGAGATGACAAAGGCATATCGCCAACTAC
LBICT7  CGAGGGAGATATATACGGATCCGGTGGGGAGTTATTTACCTACAAAGGTAATTTTATGATGTTATGTCCTGGTATGAGATGACAAAGGCATATCGCCAACTAC
LBICY4  CGAGGGAGATATATACGGATCCGGTGGGGAGTTATTTACCTACAAAGGTAATTTTATGATGTTATGTCCTGGTATGAGATGACAAAGGCATATCGCCAACTAC
LBIC-004 CGAGGGAGATATATACGGATCCGGTGGGGAGTTATTTACCTACAAAGGTAATTTTATGATGTTATGTCCTGGTATGAGATGACAAAGGCATATCGCCAACTAC

```

Figura 5. Alineamiento de un fragmento de la secuencia de nucleótidos de los genes parciales obtenidos de las cepas LBICR6, LBICS9, LBICT4, LBICT7, LBICY4, LBIC-004.

Tabla 2. Cepas analizadas que mostraron un porcentaje de identidad con genes *cry*.

Cepas	Especie	Posible Gen	% identidad (gen parcial)	Blast Proteína
LBIS14-1, LBIS14-2, LBIS15, LBIS21, LBI167, LBIS222.	<i>Lysinibacillus sphaericus</i>	-----	-----	Proteína hipotética
IR	<i>Bacillus</i> sp.	-----	-----	Policétido sintasa
LBICR6, LBICS9, LBICT4, LBICT7, LBICY4, LBIC-004.	<i>Bacillus cereus</i>	Cry8	65%	Proteína Cry8Ca

7.1 Identificación y caracterización de la cepa LBIC-004

Los genes utilizados para la caracterización e identificación de la cepa LBIC-004 fueron el gen *16S rRNA*, el gen *gyrB* y el gen de la flagelina (*hag*), los cuales fueron amplificados (Figura 6) y secuenciados. La secuencia obtenida del gen *16S rRNA* fue analizada mediante el programa Blastn y mostró un 99% de identidad con cepas bacterianas que pertenecen al grupo de *B. cereus*. Incluyendo *B. cereus*, *B. thuringiensis*, *B. mycoides*, *B. weihenstephanensis* y *B. pseudomycooides*. Sin embargo, esta secuencia tuvo un Max score con seis cepas de *B. cereus* y no así para *B. thuringiensis* o cualquier otra bacteria perteneciente al grupo *B. cereus*. Las secuencias del gen *gyrB* traducida a aminoácidos confirmó

el mayor porcentaje de similitud para *B. cereus* con un 100% de identidad. La secuencia de nucleótidos de la flagelina mostró un 98% de identidad con *B. thuringiensis* serovar entomocidus y un 97% con *B. cereus*. En cuanto a la secuencia traducida a aminoácidos de la flagelina mostró un 96% para *B. cereus* y *B. thuringiensis* serovar entomocidus. Por otra parte, la cepa LBIC-004 no es capaz de formar inclusiones paraesporales durante la fase de esporulación. Teniendo en cuenta esto y los resultados obtenidos del análisis de los genes *16S rRNA*, *gyrB* y *hag* y comparando con lo descrito y reportado por Ruiu y cols. (2015) podemos considerar a esta cepa LBIC-004 como *B. cereus*.

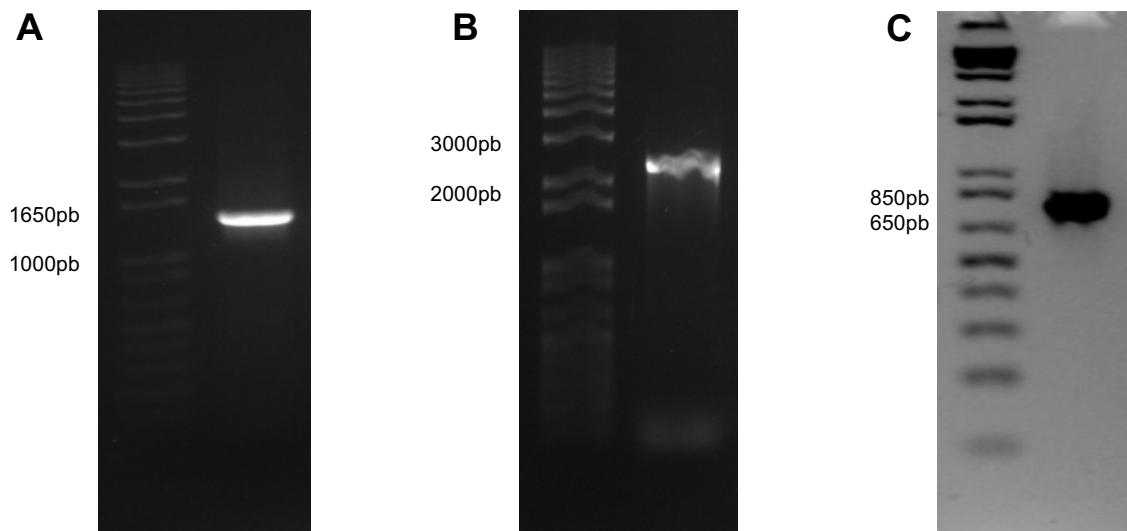


Figura 6. Amplificación de los genes *16S rRNA*, *gyrB* y *hag*. (A) Gel de agarosa mostrando el amplicon obtenido del gen *16S rRNA*. (B) Gel de agarosa mostrando el amplicon obtenido del gen *gyrB*. (C) Gel de agarosa mostrando el amplicon obtenido del gen *hag*.

7.2 Patrón de plásmidos

B. thuringiensis regularmente puede contener múltiples plásmidos; sin embargo, en general todos los patrones de plásmidos son únicos para cada cepa, por esta razón los patrones de plásmidos son herramientas valiosas para caracterizar las cepas de *B. thuringiensis*, así como cepas de *B. cereus* (Reyes-Ramírez e Ibarra 2008). La Figura 7 muestra un patrón de plásmidos de la cepa

LBIC-004, este patrón muestra seis plásmidos pequeños con tamaños que oscilan desde >12 kpb a 3 kpb; además se puede observar que la cepa contiene dos mega plásmidos. Cabe mencionar que se realizó el patrón de plásmidos para las 6 cepas que dieron positivo para genes *cry* y todas mostraron el mismo patrón de plásmidos (Anexo 1), esto junto con los resultados obtenidos del alineamiento del gen *16S rRNA*, *gyrB* y la similitud de las secuencias parciales de los genes *cry* de las cepas LBIC-R6, LBIC-S9, LBIC-T4, LBIC-T7, LBIC-Y4, LBIC-004 confirmó que estas cepas son replicas del mismo aislado. Por otra parte, como se mencionó anteriormente, el gen de la flagelina (*hag*) de la cepa LBIC-004 tuvo un 96% de identidad con el gen de la flagelina del serovar entomocidus; sin embargo, Reyes-Ramírez e Ibarra (2008) reportaron el patrón de plásmidos del serovar entomocidus, el cual carece de plásmidos pequeños y solo cuenta con un mega plásmido; por lo tanto, el patrón de plásmidos junto con el análisis de los genes *16s rRNA*, *gyrB* y *hag* nos indica que efectivamente la cepa LBIC-004 pertenece a la especie *B. cereus*.

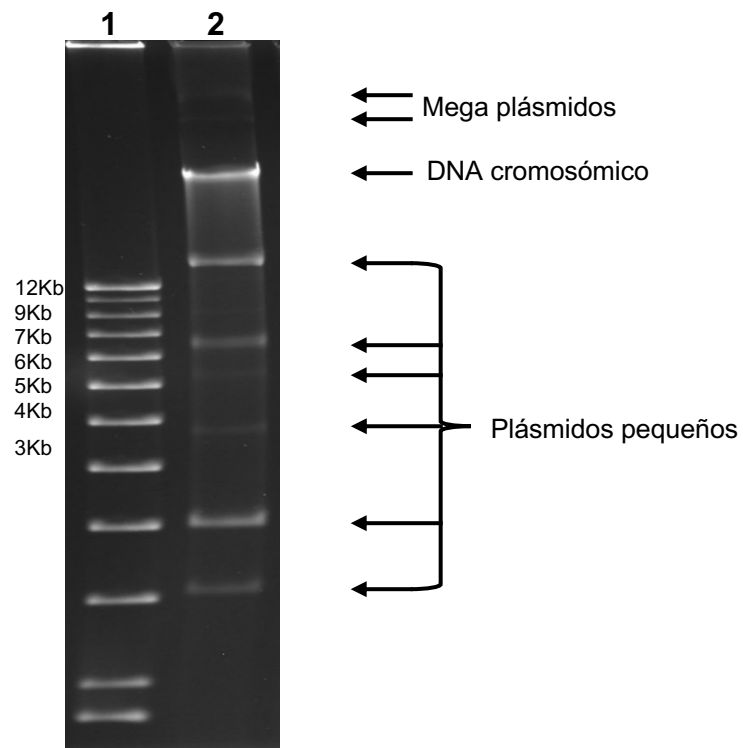


Figura 7. Patrón de plásmidos de la cepa LBIC-004. Gel de agarosa mostrando el patrón de plásmidos de *B. cereus* cepa LBIC-004 (Carril 2). Carril 1: 1kb Plus DNA Ladder.

7.3 Identificación, clonación, análisis y clasificación del gen *cry*

Utilizando los iniciadores universales reportados por Noguera e Ibarra (2010) se logró obtener un producto de amplificación de la cepa LBIC-004 de un tamaño de 1300 pb (Figura 8A) el cual fue clonado en el vector pCR 4-Topo® (Invitrogen). Las clonas transformadas fueron analizadas con varias enzimas de restricción, sin embargo, no se observó variación en los patrones de restricción (Anexo 2), lo cual indicó que el amplicón clonado se originó a partir de un único gen *cry* contenido en la cepa LBIC-004. Este amplicón fue secuenciado y el análisis de la secuencia parcial del gen indicó un máximo de 65% de identidad con el gen *cry8Ca1*. En base a este resultado y a la similitud con los genes *cry8* se intentó amplificar el gen completo utilizando iniciadores para la familia de los genes *cry8* (Tabla 1); sin embargo, no se logró obtener ningún amplicón. A partir de este resultado se utilizó el Genome Walker universal kit (Clontech) para obtener los amplicones faltantes de la secuencia completa del gen.

Una vez obtenidos los amplicones restantes, tanto río arriba y río debajo de la secuencia conocida del gen, éstos fueron secuenciados, ensamblados en una secuencia consenso y el gen completo fue analizado mediante BLASTp (Anexo 3), mostrando una identidad máxima del 41% con la secuencia del gen *cry8Ca1*, lo cual indica que se trata de un gen *cry* totalmente nuevo con un tamaño de 1953 pb (Figura 8B) que codifica para una proteína de 651 aminoácidos con un peso molecular calculado de 74.9 kDa (número de acceso del GenBank MF630935).

Por otra parte, se logró la clonación del gen *cry-like* completo en el vector pJET1.2/blunt (Thermo Fisher) y se transformó en la cepa DH5 α . Se comprobó la correcta inserción del gen en el vector con las enzimas *EcoRI*, *Sall* y *SphI*, además se comprobó por amplificación con PCR (Figura 8C). Por último, se obtuvo la clonación del gen *cry-like* en el vector de expresión pSTAB utilizando los sitios *Sall* y *SphI* y se obtuvo la transformante en DH5 α . Se corroboró la clonación

correcta del gen *cry-like* en el vector pSTAB con la enzima *EcoRI*, una doble digestión con *Sall/SphI*, por PCR y secuenciación (Figura 8C).

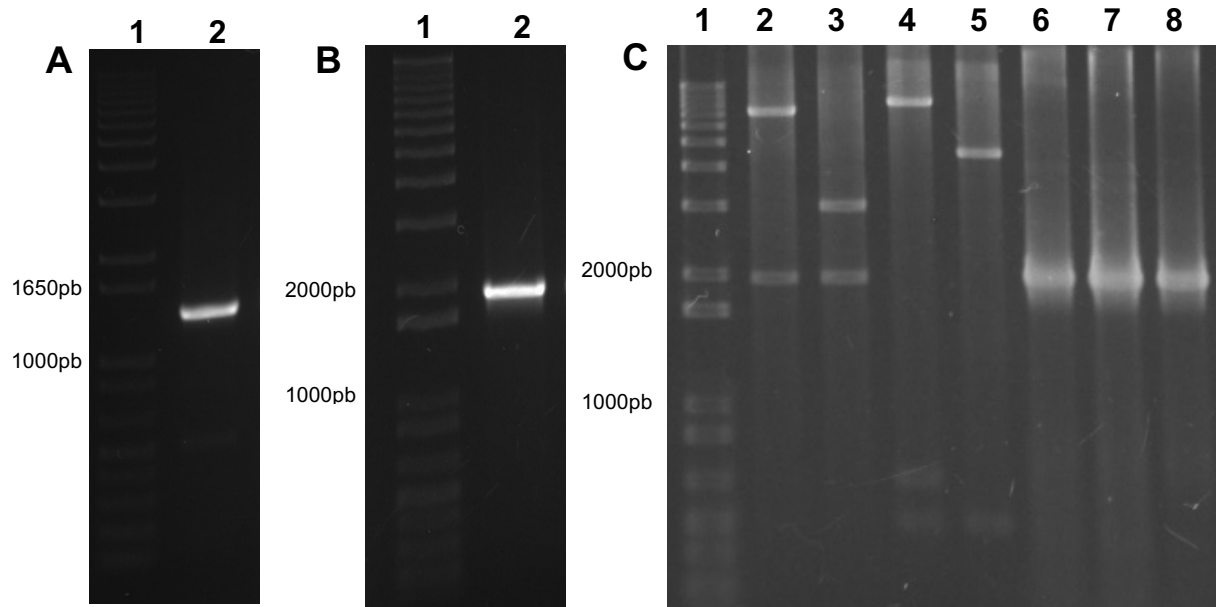


Figura 8. Amplificación parcial del gen *cry-like* y amplificación completa y clonación del gen *cry-like*. (A) Gel de agarosa mostrando el amplicon parcial (1300pb) obtenido del gen *cry* con los iniciadores Bloque 1 y Bloque 5 (Carril 2). Carril 1 MWM (1 kb Plus DNA Ladder). (B) Gel de agarosa mostrando el gen *cry-like* completo (1953pb) de la cepa LBIC-004 (Carril 2). Carril 1 MWM (1 kb Plus DNA Ladder). (C) Gel de agarosa mostrando la clonación del gen *cry-like* completo en el vector pJET1.2/blunt y en el vector de expresión pSTAB. Carril 1: MWM (1 kb Plus DNA Ladder), Carril 2: pStab:cry-like/*Sall* y *SphI*, Carril 3: pJet:cry-like/ *Sall* y *SphI*, Carril 4: pStab:cry-like/*EcoRI*, Carril 5: pJet:cry-like/*EcoRI*, Carril 6: pStab:cry-like amplificado por PCR, Carril 7: pJet:cry-like amplificado por PCR, Carril 8: LIC-004-Cry amplificado por PCR (Ctrl +).

Por otro lado, mediante el software Phyre2 se generó un modelo 3D (Figura 9) de la proteína Cry-like, donde el 96% de los residuos fueron modelados con un 100% de confianza. El modelo presenta una estructura tridimensional predicha conformada por tres dominios con un 39% de identidad con la proteína Cry8Ea1. El dominio I está conformado por α -hélices y el dominio II y el dominio III formado por laminas- β .



Figura 9. Estructura terciaria de la proteína Cry-like. Modelo 3D de la proteína Cry-like predicho con un 100% de confianza y un 39% de identidad con la proteína Cry8Ea1.

7.4 Expresión y análisis de perfil de proteínas (SDS-PAGE)

Se logró transformar cepas acristalíferas de *Bt* con el vector pSTAB-*cry-like*, las cuales se cultivaron hasta el estado de esporangio y se observaron al microscopio mediante microscopia de contraste de fases. Sin embargo, ninguna cepa transformada (4Q7 y Cry^B) fue capaz de formar inclusiones proteicas (Figura 10), la transformación correcta de las cepas fue corroborada por PCR y secuenciación.

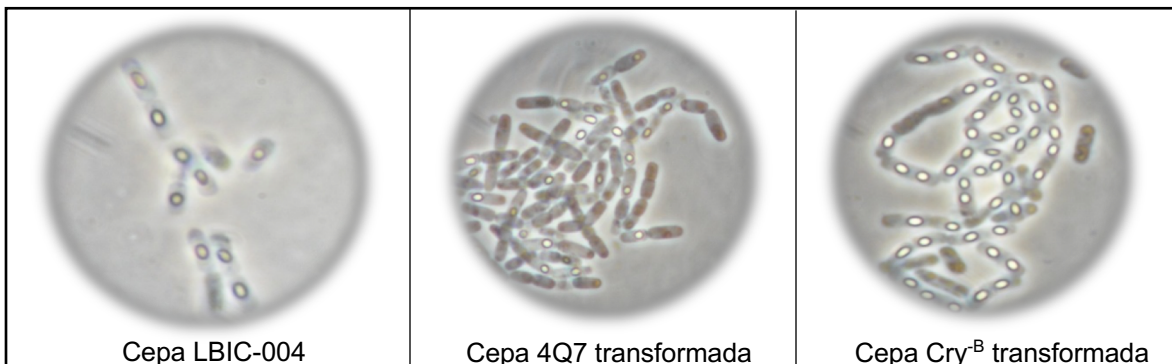


Figura 10. Estado de esporangio de la cepa nativa y las cepas recombinantes de *B. thuringiensis*.

El análisis del perfil proteico de la cepa LBIC-004 y de la cepa 4Q7-cry-like fue analizado por SDS-PAGE (Figura 11). La cepa LBIC-004 mostró una banda del tamaño esperado de 74.9 kDa, pero en la cepa 4Q7-cry-like no se observó la banda del tamaño esperado; sin embargo, hay una banda sobre-expresada de 25 kDa. Cada una de las bandas se escindió del gel y se sometió a digestión con tripsina (Hernández-Soto *et al.*, 2009). La identificación de estas bandas se logró por análisis MALDI-TOF MS. La banda de 74.9 kDa de la cepa LBIC-004 y la banda de 25 kDa de la cepa 4Q7-cry-like mostraron una correspondencia con la secuencia de aminoácidos de la proteína Cry-like (Anexo 4). Con estos resultados, se deduce que, a pesar de que la cepa LBIC-004 es capaz de expresar la proteína Cry-like, ésta no tiene la capacidad de formar cristales durante la fase de esporulación. Esto puede deberse a que la proteína carece de la mitad C-terminal. Se ha observado que la mitad C-terminal es esencial en la expresión y cristalización de diversas proteínas Cry (Barboza-Corona *et al.*, 2012 y Sun *et al.*, 2013).

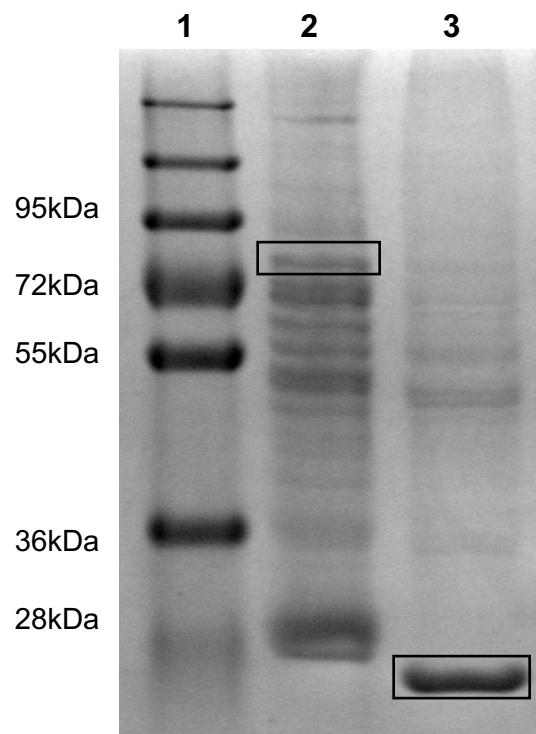


Figura 11. Perfil proteico de la cepa LBIC-004 y 4Q7 recombinante. Gel SDS-PAGE mostrando el perfil proteico de la cepa LBIC-004 y 4Q7-cry-like. Carril 1: MWM (PageRuler Protein Ladder). Carril 2: Expresión de la proteína Cry-like en la cepa LBIC-004 con un peso de 74.9 kDa. Carril 3: Expresión de la proteína Cry-like en la cepa 4Q7-cry-like con un peso de 25 kDa.

7.5 Bioensayos

Para tratar de detectar alguna actividad tóxica tanto de la cepa nativa LBIC-004 como de la recombinante 4Q7-cry-like se realizaron bioensayos contra larvas de insectos y nemátodos. Estos incluyeron *A. aegypti*, *M. sexta* y *Phyllophaga* sp., así como *C. elegans*. A pesar de las altas concentraciones utilizadas, no se observó mortalidad en ninguna de las especies de insectos. Por otro lado, se logró observar un efecto en la disminución de la población de nemátodos, al ser alimentados con la cepa LBIC-004 en estado de esporangio comparándola con el control negativo OP50 y *Bt kurstaki*; sin embargo, este resultado no es significativo (Anexo 5).

7.6 Agrupamiento de proteínas Cry obtenidos mediante el uso de CD-HIT y construcción del perfil HMM

Hasta enero del 2017 habían reportadas 791 secuencias de genes *cry*. Sin embargo, algunas secuencias se encontraban incompletas, por lo cual se decidió trabajar sólo con las secuencias que se encuentran completas como se menciona en la sección de Materiales y Métodos. Se obtuvieron alineamientos desde un 50% hasta un 90% de similitud; y como era de esperarse, cuando se hacían agrupamientos a un porcentaje de identidad bajo como 50% se obtuvieron menos agrupamientos, pero cada grupo tenía mayor número de secuencias; por el contrario, a un porcentaje alto de identidad como 90% se obtuvieron más grupos, pero con menor número de secuencias (datos no mostrados).

Al analizar los alineamientos de las secuencias de las proteínas Cry en estos clústers observamos claramente que la región amino de las secuencias ya no es tan conservada debido probablemente al incremento de secuencias de las proteínas Cry descubiertas, sin embargo, el extremo carboxilo terminal de las proteínas Cry está altamente conservado. A partir de estas observaciones se decidió hacer el perfil HMM a partir de la mitad carboxilo de las proteínas Cry, por

lo tanto, se seleccionaron todas las secuencias de las proteínas Cry que tuvieran más de 800 aa, es decir aquellas proteínas Cry de tres dominios que no fueran naturalmente truncadas y estas se agruparon con un 70% de similitud incluyendo los criterios ya mencionados. En total se obtuvieron 6 clústeres con un total de 111 secuencias, las cuales se “podaron” manualmente eliminando los primeros 800 aa, quedando solo el extremo carboxilo terminal de las proteínas Cry; éstos se alinearon y a partir de este alineamiento de las 111 secuencias de los carboxilos terminales se construyó el perfil HMM (Figura 12) con el paquete de software HMMER versión 3.1b2 con un total de 586 posiciones consenso.

7.7 Análisis del plásmido pBtoxis y Cepa 4Q7 con el perfil HMM

Una vez diseñado el perfil HMM se validó utilizando las secuencias proteicas del genoma de *B. thuringiensis israelensis* (*Bti*) y de *B. thuringiensis* serovar *israelensis* cepa 4Q7 para probar que realmente estuviera funcionando y no arrojara falsos positivos o falsos negativos.

La cepa *Bti* contiene un mega-plásmido llamado pBtoxis el cual contiene las toxinas Cry4Aa, Cry4Ba, Cry10Aa y Cry11Aa, por lo tanto, serviría como control positivo al detectar solamente las proteínas Cry4Aa y Cry4Ba que son las únicas proteínas Cry en este plásmido que tienen su extremo carboxilo terminal. El modelo HMM fue capaz de identificar estas dos proteínas (Cry4Aa y Cry4Ba). Además de estas dos proteínas, el modelo identificó 3 anotaciones más, identificadas como pBt025, pBt026 y pBt048. Las dos primeras corresponden a fragmentos de proteínas Cry28-like y el pBt048 es similar al extremo carboxilo de la proteína Cry4. Como control negativo del modelo HMM se utilizó el genoma de la cepa 4Q7 que se encuentra curada del plásmido y por lo tanto carece de genes *cry*; en este caso el modelo no detecto ningún alineamiento.

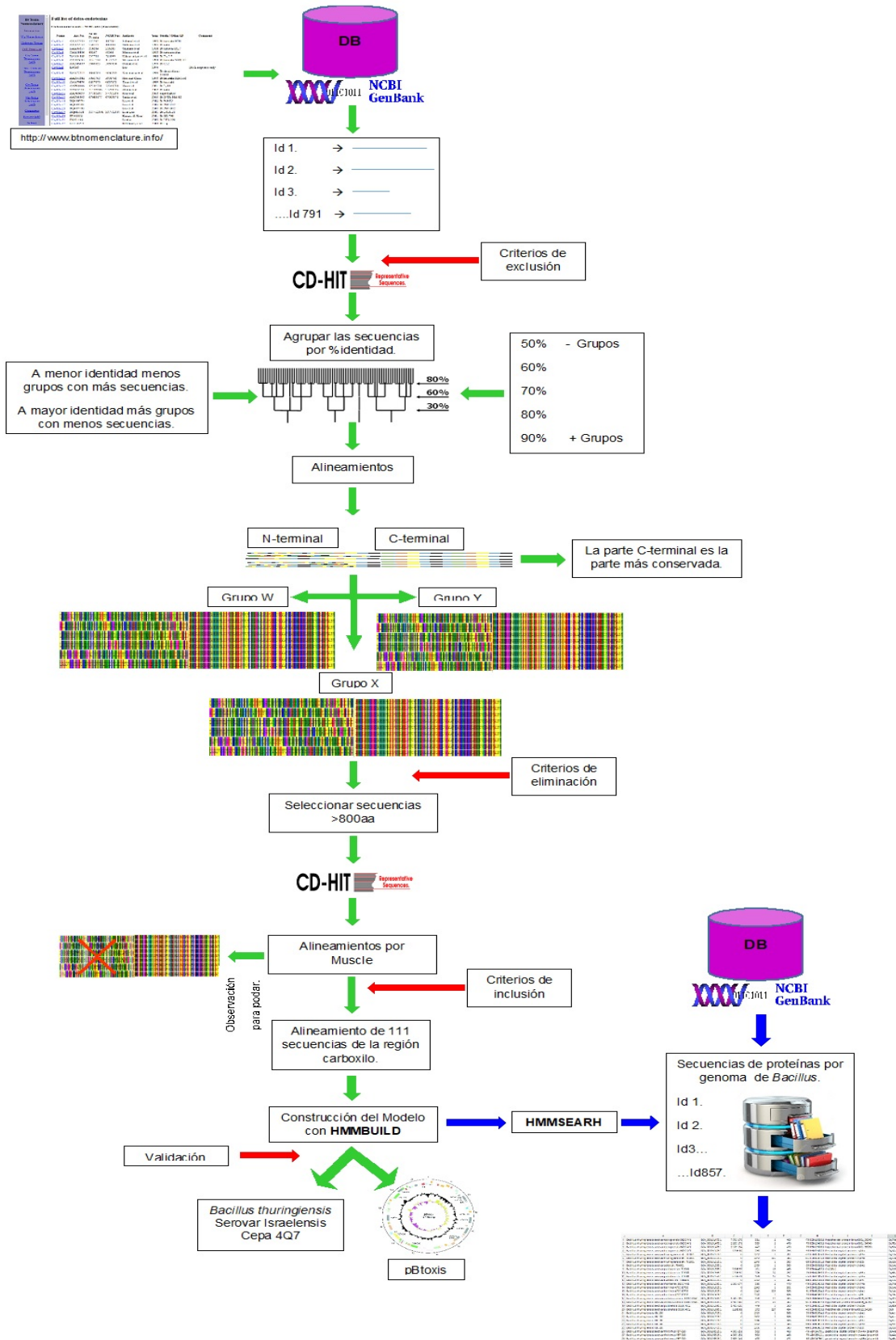


Figura 12. Diagrama para construcción del perfil HMM para búsqueda de proteínas Cry.

7.8 Identificación de proteínas Cry mediante *hmmsearch* en genomas de *Bacillus* spp.

Como ya se mencionó en la sección de Materiales y Métodos, se realizó una base de datos de todos los genomas de *Bacillus* incluyendo a *Bt*, esto con la finalidad de identificar los genes ya reportados y aquellos que aún no han sido anotados como genes *cry*. En total se descargaron 857 genomas de *Bacillus* con los cuales fue probado el modelo HMM, el cual fue capaz de identificar un total de 174 proteínas (Anexo 6) distribuidas principalmente en cuatro especies de *Bacillus* (*B. thuringiensis*, *B. cereus*, *B. subtilis*, *Bacillus* sp cepa L1B05 y *Bacillus* sp cepa L1B08). De las 174 proteínas que identificó el modelo HMM solo 125 proteínas fueron identificadas como proteínas Cry, 44 como proteínas hipotéticas y 5 como productos proteicos sin nombre (Tabla 3).

Tabla 3. Resultados obtenidos utilizando el perfil HMM.

Principales datos obtenidos con el perfil HMM	
Genomas analizados	857 genomas de <i>Bacillus</i> (Incluyendo Bt)
Proteínas identificadas	174 Proteínas
Especies de <i>Bacillus</i>	<i>B. thuringiensis</i> , <i>B. cereus</i> , <i>B. subtilis</i> , <i>Bacillus</i> sp cepa L1B05 y <i>Bacillus</i> sp cepa L1B08
Proteínas Cry	125 Proteínas Cry
Proteínas hipotéticas	44 (Proteínas Cry)
Productos proteicos sin nombre	5 (Proteínas Cry)
Tipos de proteínas Cry encontrados	Cry1, Cry4, Cry5, Cry7, Cry8, Cry9, Cry14, Cry21, Cry32, Cry39, Cry40 y Cry61

Las 44 secuencias de las proteínas hipotéticas y las 5 secuencias de los productos proteicos fueron analizadas manualmente usando la herramienta Blastp del NCBI y todas tuvieron una coincidencia con proteínas Cry.

En la Figura 13 se muestra los 12 grupos de proteínas Cry que se identificaron en los genomas mediante el perfil HMM, así como el número de secuencias de proteínas Cry encontrado y el porcentaje de cada grupo. Como era de esperar, las proteínas tipo Cry1 se encontraron con mayor frecuencia (43.68%) representadas por 15 subgrupos diferentes, seguidas por el tipo Cry4 (19.54%), representado por cuatro subgrupos. En la Tabla 4 se muestran los subgrupos encontrados para cada tipo de proteínas Cry.

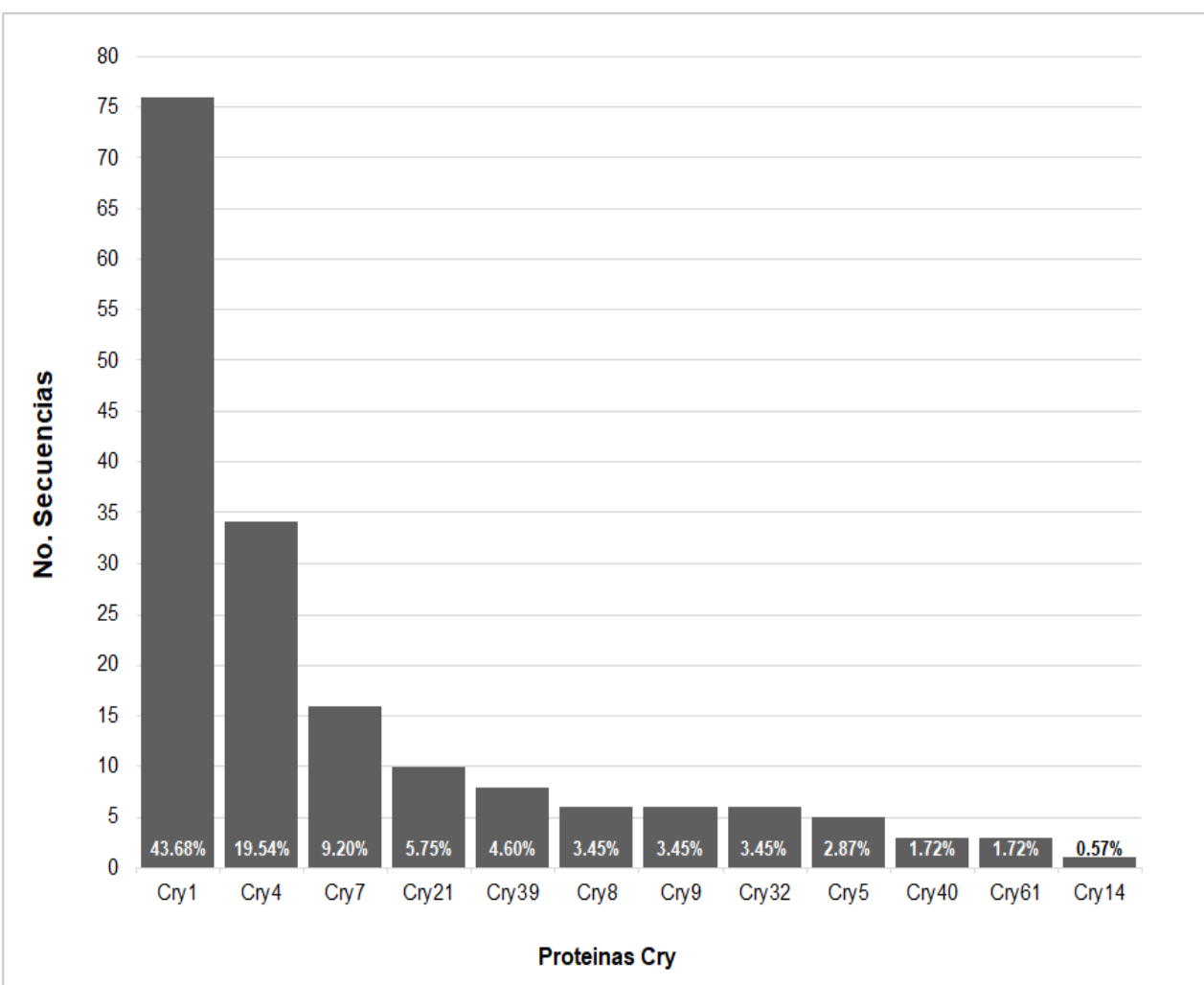


Figura 13. Distribución de las proteínas Cry obtenidas del análisis *in silico* de 857 genomas de *Bacillus* spp. mediante el perfil HMM.

Tabla 4. Principales subgrupos de proteínas Cry identificadas *in silico* por el modelo HMM en 857 genomas de *Bacillus* spp.

Proteínas Cry	No. Secuencias	Subgrupos	Frecuencia (%)
Cry1	76	Cry1Aa, Cry1Ab, Cry1Ac, Cry1Ae, Cry1Ba, Cry1Bb, Cry1Bc, Cry1Bd, Cry1Ca, Cry1Cb, Cry1Da, Cry1Db, Cry1Fb, Cry1Ga, Cry1Ha	43.68
Cry4	34	Cry4Aa, Cry4Ba, Cry4Cb, Cry4Da	19.54
Cry7	16	Cry7Aa, Cry7Ab, Cry7Ba, Cry7Ca, Cry7Ea	9.20
Cry21	10	Cry21Ba, Cry21Fa, Cry21Ga, Cry21Ha	5.75
Cry39	8	Cry39	4.60
Cry8	6	Cry8Aa, Cry8Ba, Cry8Ea, Cry8X	3.45
Cry9	6	Cry9Ea	3.45
Cry32	6	Cry32Da, Cry32Eb	3.45
Cry5	5	Cry5Aa, Cry5Ba	2.87
Cry40	3	Cry40Aa	1.72
Cry 61	3	Cry61Aa	1.72
Cry14	1	Cry14Aa	0.57

Sin embargo, lo más importante de las 147 proteínas identificadas mediante el modelo HMM son las 42 proteínas que se encontraron en especies diferentes a *B. thuringiensis*. De estas proteínas 36 se encontraron en genomas de *B. cereus*, cinco en bacilos no identificados (*Bacillus* spp.) y una en un genoma de *B. subtilis*. De forma similar al recuento total, las proteínas de tipo Cry1 fueron más frecuentes en *B. cereus*, entre otros tipos. Sólo Cry4A y Cry40A se encontraron en genomas de bacilos no identificados y se encontró una Cry14Aa en un genoma de *B. subtilis* (Tabla 5).

Tabla 5. Subgrupos de proteínas Cry encontrados en especies diferentes a *B. thuringiensis* mediante el modelo HMM.

Especie	Secuencias encontradas	Subgrupos de proteínas Cry encontradas	Secuencias por grupo	Identidad (%)
<i>Bacillus cereus</i>	36	Cry1Aa	3	100*
		Cry1Ab	2	100*
		Cry1Ac	4	100*
		Cry1Ca	1	99
		Cry1Da	1	100
		Cry4Ba	11	63 - 100
		Cry5Aa	2	100*
		Cry8Aa	2	88, 90
		Cry9Ea	2	100*
		Cry21Ba	2	34, 46
		Cry21Fa	2	43, 49
		Cry32Aa	1	56
		Cry32Da	2	59, 74
		Cry32Eb	1	73
<i>Bacillus subtilis</i>	1	Cry14Aa	1	100
<i>Bacillus</i> spp.	5	Cry4Aa	3	99 – 100
		Cry40	2	62*

*Todas las secuencias muestran el mismo nivel de identidad.

7.9 Análisis MLST y reconstrucción filogenética

Sin embargo, está bien establecido que las cepas de *B. thuringiensis* pueden registrarse erróneamente como *B. cereus*, además algunos reportes recientes indican que probablemente algunas cepas del grupo *cereus* se encuentran mal identificadas en el GenBank. Por lo tanto, se realizó un análisis detallado de MLST para la autenticación de los genomas obtenidos del GenBank y utilizados en este estudio (Tabla 6); en términos generales el análisis de MLST es razonablemente preciso para determinar la especie.

Tabla 6. Identificación por MLST de las bacterias cuyos genomas mostraron secuencias Cry, y su comparación con las especies registradas en el GenBank, usadas para el análisis *in silico* de este estudio y el perfil alélico encontrado para cada cepa.

No. Acceso	Designación original	Especie propuesta	Perfil alélico							ST
			<i>glp</i>	<i>gmk</i>	<i>ilv</i>	<i>pta</i>	<i>pur</i>	<i>pyc</i>	<i>tpi</i>	
GCA_000344745.1	<i>Bacillus subtilis</i>	Type strain	-	-	-	-	-	-	-	1
GCF_000299035.1	" <i>Bacillus bingmayongensis</i> "	Type strain	193	114	221	199	189	168	154	763
GCF_000496285.1	<i>Bacillus toyonensis</i>	Type strain	43	26	35	42	39	41	30	111
GCF_000017425.1	<i>Bacillus cytotoxicus</i>	Type strain	211	127	221	221	57	10	171	930
GCF_000712595.1	" <i>Bacillus manliponensis</i> "	Type strain	205	123	107	213	199	10	165	907
GCA_000712615.1	" <i>Bacillus gaemokensis</i> "	Type strain	195	121	230	211	197	177	161	768
GCA_000161455.1	<i>Bacillus pseudomycooides</i>	Type strain	63	13	58	23	25	44	35	83
GCA_000007825.1	<i>Bacillus cereus</i> ATCC 14579	Type strain	13	125	8	11	11	12	169	921
GCA_000513155.1	<i>Bacillus weihenstephanensis</i>	Type strain	64	10	79	36	56	22	11	196
GCA_000003925.1	<i>Bacillus mycooides</i>	Type strain	25	10	22	53	57	23	44	116
GCA_000181835.1	<i>Bacillus anthracis</i>	Type strain	1	1	1	1	1	1	1	1
GCA_001640965.1	<i>Bacillus thuringiensis</i> serovar alesti	<i>B. thuringiensis</i>	15	7	7	2	7	10	13	12
GCA_001598095.1	<i>Bacillus thuringiensis</i>	<i>B. thuringiensis</i>	15	7	7	2	5	8	13	23
GCA_001595725.1	<i>Bacillus thuringiensis</i>	<i>B. cereus</i>	232	8	48	45	58	87	7	1133
GCA_001583975.1	<i>Bacillus cereus</i>	<i>B. anthracis</i>	74	137	85	62	66	66	186	1085
GCA_001583935.1	<i>Bacillus cereus</i>	<i>B. cereus</i>	22	9	18	18	20	19	7	33
GCA_001548175.1	<i>Bacillus thuringiensis</i> serovar tolworthi	<i>B. cereus</i>	12	8	8	11	9	12	23	22
GCA_001420855.1	<i>Bacillus thuringiensis</i>	<i>B. cereus</i>	7	8	16	13	2	16	7	8
GCA_001402735.1	<i>Bacillus thuringiensis</i>	<i>B. thuringiensis</i>	15	7	7	2	6	8	13	16
GCA_001286785.1	<i>Bacillus subtilis</i>	<i>B. cereus</i> *	56	1	93	93	51	37	71	ND
GCA_001183785.1	<i>Bacillus thuringiensis</i> serovar indiana	<i>B. cereus</i>	37	9	14	87	12	14	7	241
GCA_001182785.1	<i>Bacillus thuringiensis</i>	<i>B. cereus</i> *	126	152	76	46	4	114	37	ND
GCA_001044855.1	<i>Bacillus cereus</i>	<i>B. cereus</i>	9	8	16	13	2	16	9	15
GCA_001017635.1	<i>Bacillus thuringiensis</i>	<i>B. cereus</i>	7	8	16	13	2	16	7	8
GCA_001008665.1	<i>Bacillus cereus</i>	<i>B. cereus</i>	9	8	16	13	2	16	9	15
GCA_000969665.1	<i>Bacillus thuringiensis</i> serovar mexicanensis	<i>B. cereus</i>	7	8	16	13	2	16	7	8
GCA_000878525.1	<i>Bacillus thuringiensis</i> Sbt003	<i>B. cereus</i>	12	8	11	11	9	13	5	17
GCA_000836745.1	<i>Bacillus</i> sp. L_1B0_8	<i>B. thuringiensis</i>	126	146	133	46	4	114	37	1161
GCA_000836715.1	<i>Bacillus</i> sp. L_1B0_5	<i>B. thuringiensis</i>	126	146	133	46	4	114	37	1161
GCA_000835235.1	<i>Bacillus thuringiensis</i> serovar kurstaki	<i>B. cereus</i>	7	8	16	13	2	16	7	8
GCA_000833655.1	<i>Bacillus thuringiensis</i> serovar morrisoni	<i>B. thuringiensis</i>	15	7	7	2	5	8	13	23
GCA_000803665.1	<i>Bacillus thuringiensis</i> serovar galleriae	<i>B. cereus</i>	9	8	16	13	2	16	9	15
GCA_000774075.2	<i>Bacillus thuringiensis</i>	<i>B. cereus</i> *	24	12	33	16	48	31	19	ND
GCA_000747545.1	<i>Bacillus thuringiensis</i> serovar kurstaki str. YBT-1520	<i>B. cereus</i>	7	8	16	13	2	16	7	8
GCA_000717535.1	<i>Bacillus thuringiensis</i> serovar kurstaki str. HD-1	<i>B. cereus</i> *	7	8	16	13	2	16	7	ND

GCA_000710255.1	<i>Bacillus thuringiensis</i> serovar kurstaki str. HD-1	<i>B. cereus</i>	7	8	16	13	2	16	7	8
GCA_000688795.1	<i>Bacillus thuringiensis</i> serovar kurstaki str. YBT-1520	<i>B. cereus</i> *	7	8	16	13	2	16	169	ND
GCA_000600315.1	<i>Bacillus thuringiensis</i>	<i>B. cereus</i>	30	8	16	13	2	16	7	25
GCA_000530375.1	<i>Bacillus thuringiensis</i> DB27	<i>B. cereus</i>	7	9	152	12	12	14	7	545
GCA_000497525.2	<i>Bacillus thuringiensis</i> YBT-1518	<i>B. thuringiensis</i>	15	143	10	242	4	7	14	1261
GCA_000468995.3	<i>Bacillus thuringiensis</i> serovar aizawai str. Hu4-2	<i>B. cereus</i>	9	8	16	13	2	16	9	15
GCA_000468975.2	<i>Bacillus thuringiensis</i> serovar aizawai str. Leapi01	<i>B. cereus</i>	9	8	16	13	2	16	9	15
GCA_000399505.1	<i>Bacillus cereus</i> ISP2954	<i>B. cereus</i>	7	8	16	13	2	16	7	8
GCA_000399365.1	<i>Bacillus cereus</i> HuB13-1	<i>B. cereus</i>	7	8	16	13	2	16	7	8
GCA_000399345.1	<i>Bacillus cereus</i> BMG1.7	<i>B. cereus</i>	7	8	16	13	2	16	7	8
GCA_000399265.1	<i>Bacillus cereus</i> BAG20-1	<i>B. thuringiensis</i>	126	120	228	209	4	114	37	783
GCA_000399105.1	<i>Bacillus cereus</i> BAG1X2-3	<i>B. thuringiensis</i>	126	120	228	209	4	114	37	783
GCA_000399085.1	<i>Bacillus cereus</i> BAG1X2-2	<i>B. thuringiensis</i>	126	120	228	209	4	114	37	783
GCA_000399065.1	<i>Bacillus cereus</i> BAG1X2-1	<i>B. thuringiensis</i>	126	120	228	209	4	114	37	783
GCA_000399005.1	<i>Bacillus cereus</i> BAG1X1-1	<i>B. thuringiensis</i>	126	120	228	209	4	114	37	783
GCA_000387405.2	<i>Bacillus thuringiensis</i> T01-328	<i>B. thuringiensis</i>	15	6	10	8	3	7	14	10
GCA_000341665.1	<i>Bacillus thuringiensis</i> serovar thuringiensis str. IS5056	<i>B. thuringiensis</i>	15	6	10	8	3	7	14	10
GCA_000338755.1	<i>Bacillus thuringiensis</i> serovar kurstaki str. HD73	<i>B. cereus</i>	7	8	16	13	2	16	7	8
GCA_000300475.1	<i>Bacillus thuringiensis</i> MC28	<i>B. toyonensis</i>	72	42	69	42	63	41	30	158
GCA_000293685.1	<i>Bacillus cereus</i> HD73	<i>B. cereus</i>	7	8	16	13	2	16	7	8
GCA_000292705.1	<i>Bacillus thuringiensis</i> HD-789	<i>B. thuringiensis</i>	15	7	7	2	6	8	13	16
GCA_000292455.1	<i>Bacillus thuringiensis</i> HD-771	<i>B. thuringiensis</i>	15	7	7	2	7	10	13	12
GCA_000290795.1	<i>Bacillus cereus</i> VD154	<i>B. cereus</i>	12	8	11	11	9	13	5	17
GCA_000290695.1	<i>Bacillus cereus</i> VDM022	<i>B. mycoides</i>	64	10	22	15	83	23	11	222
GCA_000193355.1	<i>Bacillus thuringiensis</i> serovar chinensis CT-43	<i>B. thuringiensis</i>	15	6	10	8	3	7	14	10
GCA_000190515.1	<i>Bacillus thuringiensis</i> serovar finitimus YBT-020	<i>B. anthracis</i>	24	2	74	5	32	3	22	171
GCA_000161715.1	<i>Bacillus thuringiensis</i> IBL 200	<i>B. thuringiensis</i>	68	54	98	2	37	84	72	240
GCA_000161695.1	<i>Bacillus thuringiensis</i> serovar pulsensis BGSC 4CC1	<i>B. anthracis</i>	61	38	55	1	1	53	19	110
GCA_000161635.1	<i>Bacillus thuringiensis</i> serovar andalousiensis BGSC 4AW1	<i>B. cereus</i> *	56	1	113	86	18	53	64	ND
GCA_000161615.1	<i>Bacillus thuringiensis</i> serovar berliner ATCC 10792	Type strain	15	6	10	8	3	7	14	10
GCA_000161595.1	<i>Bacillus thuringiensis</i> serovar monterrey BGSC 4AJ1	<i>B. anthracis</i>	57	1	52	1	1	52	1	107
GCA_000161575.1	<i>Bacillus thuringiensis</i> serovar kurstaki str. T03a001	<i>B. thuringiensis</i>	68	54	98	2	37	84	72	240
GCA_000161555.1	<i>Bacillus thuringiensis</i> serovar pakistani str. T13001	<i>B. cereus</i>	12	8	11	11	9	13	5	17
GCA_000161535.1	<i>Bacillus thuringiensis</i> serovar sotto str. T04001	<i>B. cereus</i> *	15	7	7	2	7	10	13	ND
GCA_000161515.1	<i>Bacillus thuringiensis</i> serovar thuringiensis str. T01001	<i>B. thuringiensis</i>	15	6	10	8	3	7	14	10
GCA_000161475.1	<i>Bacillus thuringiensis</i> serovar tochiensis BGSC 4Y1	<i>B. anthracis</i>	24	12	33	37	48	31	19	104

ND: No determinado. Las cepas con el asterisco fueron enviadas al servidor MLST-1.8 (<https://cge.cbs.dtu.dk/services/MLST/>), debido a que no se encontró un perfil alélico exacto para esas cepas en la base de datos de pubMLST (<https://pubmlst.org/bcereus/>).

En la Tabla 6 se pueden observar todos los números de acceso de los genomas que identificó el modelo HMM y que dieron positivo para al menos una secuencia de una proteína Cry y su clasificación de acuerdo con la información obtenida del GenBank. Además se muestra el perfil alélico para cada uno de los genomas, éstos obtenidos de la base de datos del PubMLST.org/bcereus/ y del servidor MLST-1.8 del Center for Genomic Epidemiology, así como el ST obtenido para cada genoma a excepción de los genomas: GCA_001286785.1, GCA_001182785.1, GCA_000774075.2, GCA_000717535.1, GCA_000688795.1, GCA_000161635.1, GCA_000161535.1, que no tuvieron una coincidencia exacta en la base de datos del PubMLST.org/bcereus para alguno de los siete genes utilizados para el análisis de MLST, por lo tanto se enviaron al Center for Genomic Epidemiology donde obtuvimos el mejor perfil alélico y la posible especie.

Se obtuvieron las secuencias concatenadas de los 7 genes del MLST: *glp*, *gmk*, *ilv*, *pta*, *pur*, *pyc*, y *tpi* para cada uno de los genomas, a excepción de los que se enviaron al CGE, con las cuales se realizó la reconstrucción filogenética mediante el programa MEGA7 (Figura 14). Además, para la reconstrucción filogenética se incluyó la secuencia de la cepa tipo de *B. subtilis* como grupo externo; por otro lado se incluyeron todas las secuencias de las cepas tipo del grupo *cereus*: *B. anthracis*, *B. cereus*, *B. thuringiensis*, *B. mycooides*, *B. pseudomycooides*, *B. weihenstephanensis*, *B. cytotoxicus*, *B. toyonensis*, “*B. gaemokensis*”, “*B. manliponensis*” y “*B. bingmayongensis*” (Liu *et al.*, 2015) para determinar a qué posible especie pertenece cada genoma. En la Tabla 6 se muestran las sugerencias de los posibles cambios que podrían hacerse en cuanto a la clasificación de los genomas de acuerdo con su especie, estas sugerencias fueron tomadas del resultado obtenido de la reconstrucción filogenética (Figura 14) donde se agrupan los genomas en cada una de las secuencias tipo del grupo *cereus*. Cabe mencionar que con el análisis del MLST sólo 22 genomas coincidieron con la especie registrada en el NCBI y los otros 38 genomas mostraron una discrepancia de la especie reportada en la base de datos del NCBI con la identificación basada en nuestro análisis MLST. Por lo tanto, el número de

genomas con genes *cry* en especies diferentes a *B. thuringiensis* aumento de 18 a 39 genomas, así como el número de secuencias de proteínas Cry encontradas en especies diferentes a *B. thuringiensis* aumentó de 42 a 119, lo cual indica que las secuencias de proteínas Cry fueron más frecuentes en cepas diferentes a *B. thuringiensis*.

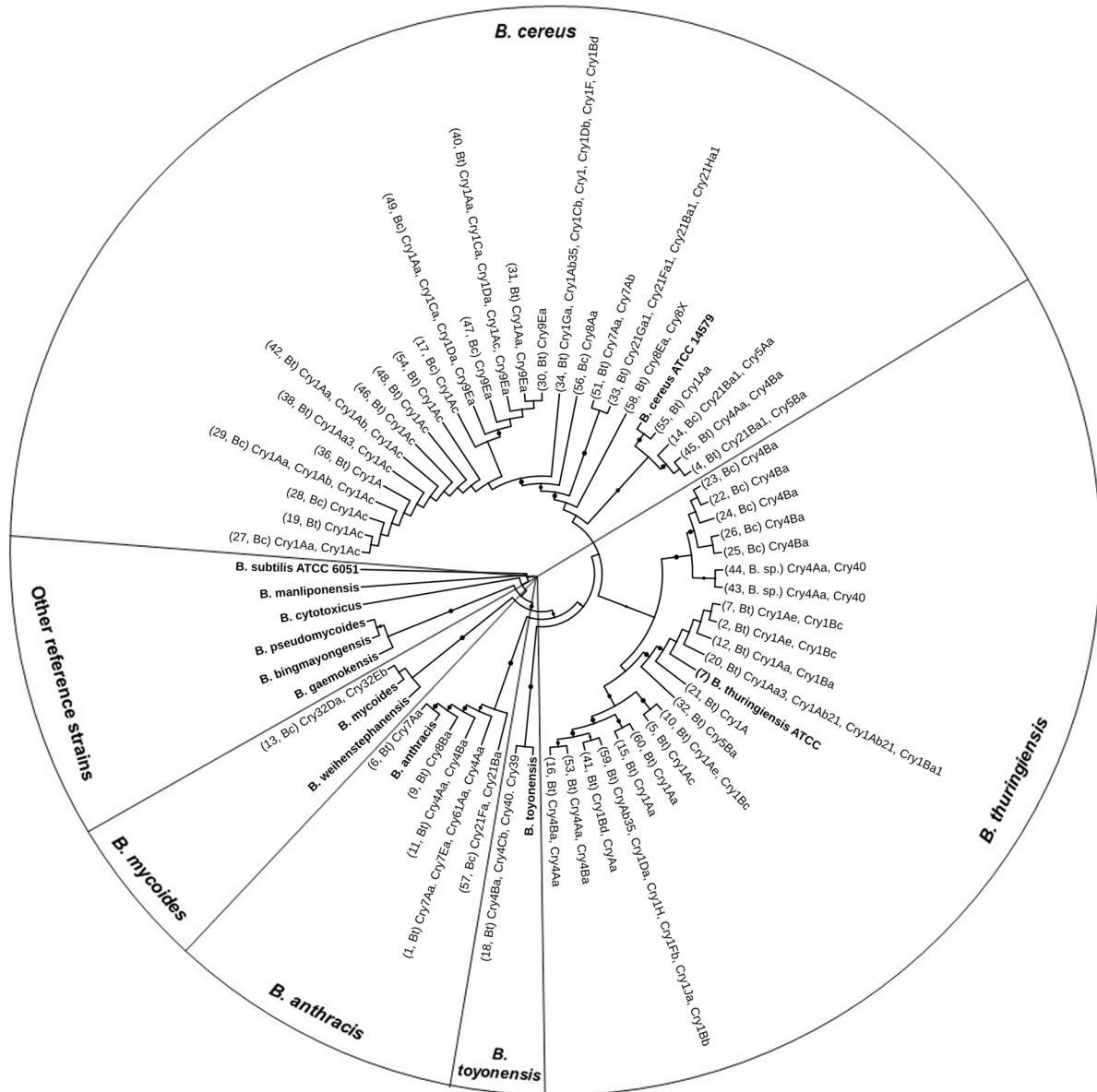


Figura 14. Árbol filogenético del análisis de MLST. Árbol inferido del MLST de 65 genomas del grupo *B. cereus* usando 7 genes endógenos. La cepa de *B. subtilis* ATCC 6051 fue utilizada como cepa tipo y como grupo externo. Los valores de Bootstrap ($> 50\%$) se muestran en los puntos de ramificación. Se muestra el árbol con la mayor probabilidad de registro (-9908.5930). Los árboles iniciales para la búsqueda heurística se obtuvieron automáticamente aplicando los algoritmos Neighbor-Join y BioNJ a una matriz de distancias por pares estimadas utilizando el enfoque Maximum Composite Likelihood (MCL) y luego seleccionando la topología con un valor de verosimilitud log superior. Se usó una distribución discreta de Gamma para modelar las diferencias de velocidad evolutivas entre los sitios (4 categorías (+ G, parámetro = 0,6120)). El análisis involucró 65 secuencias de nucleótidos. Se eliminaron todas las posiciones que contenían lagunas y datos faltantes. Hubo un total de 1960 posiciones en el conjunto de datos final. Se realizaron análisis evolutivos en MEGA7.

DISCUSIÓN

Existen muy pocos reportes sobre la identificación de genes *cry* en otras bacterias como es el caso de *C. bifermentans*, *P. popilliae*, *P. lentimorbus* y *L. sphaericus* (Barloy, *et al.*, 1996; Zhang, *et al.*, 1997; Barloy, *et al.*, 1998; Yokoyama *et al.*, 2004 y Jones *et al.*, 2007); sin embargo, a pesar de que hay muy pocos reportes, en este trabajo se planteó la idea de buscar genes *cry* en *Bacillus* diferentes a *Bt* mediante el uso de los iniciadores quasi universales reportados por Noguera e Ibarra (2010), con la finalidad de determinar que tan distribuidos se encuentran estos genes en otras bacterias y la posibilidad de encontrar nuevos genes *cry*.

Con la misma finalidad, se probó una búsqueda sistemática *in silico*, mediante el desarrollo de un perfil HMM a partir de la región C-terminal de las proteínas Cry, con el cual se realizó una búsqueda en 857 genomas relacionados con 89 especies de *Bacillus*.

Primeramente, se realizó una búsqueda exhaustiva de genes *cry* en bacterias diferentes a *Bacillus thuringiensis*; y se logró analizar un total de 223 cepas de *Bacillus* spp. donde se lograron obtener 45 amplicones, de los cuales se seleccionaron 13 amplicones de acuerdo con la probabilidad de ser genes tipo *cry*. Cabe mencionar que las 223 cepas de *Bacillus* spp. fueron seleccionadas por su morfología y que no eran capaces de formar inclusiones paraesporales durante la fase de esporulación. De los 13 amplicones seleccionados para su secuenciación solo 6 dieron un porcentaje de similitud con genes tipo *cry*; sin embargo, los 6 amplicones tenían exactamente el mismo porcentaje de similitud con el gen *cry8*, lo cual indicaba que podría tratarse del mismo gen, inclusive del mismo aislado de acuerdo con los resultados obtenidos del gen *16S rRNA* y del gen *gyrB*, por otra parte, esto puede deberse a que estas 6 cepas fueron aisladas de la laguna intermedia del sistema de agua de Churince en la cuenca de Cuatro Ciénegas ubicada en el desierto de Chihuahua (Anexo 7). Debido a que esta cepa no

presentaba la formación de cristales parasporales, fue necesario corroborar que no perteneciera a la especie *B. thuringiensis*. Para ello se llevaron a cabo varias pruebas.

Una de las características para determinar diferencias entre especies del grupo *cereus* se basa en el análisis del patrón de plásmidos en *Bacillus*, ya que son únicos para cada cepa, además es una herramienta capaz de distinguir y caracterizar las cepas de *B. thuringiensis*, así como cepas de *B. cereus* (Reyes-Ramírez e Ibarra 2008). Para diferenciar si las seis cepas que dieron positivo para los genes tipo *cry* podrían ser el mismo aislado, se realizó el patrón de plasmidos de las mismas (Anexo 1), en el cual se observó que tenían el mismo número de plásmidos y migraban exactamente igual, por lo cual solo se trabajó con una sola cepa. Sin embargo, posteriormente se realizaron otros experimentos para demostrar que se trataba del mismo aislado y el mismo gen *cry*.

Como ya se mencionó en el apartado de resultados, se utilizaron los genes *16S rRNA*, *gyrB* y *hag* para la caracterización de la cepa LBIC-004; sin embargo, antes de descartar las seis cepas positivas, todas fueron amplificadas para el gen del *16S rRNA* y el gen *gyrB*, y todas las cepas dieron positivo o con un mayor porcentaje de similitud con *Bacillus cereus* tanto con el *16S rRNA* y *gyrB*. Por otra parte, para descartar las cepas, se realizó un alineamiento de las secuencias parciales de los genes *cry* de cada cepa, el alineamiento mostró una similitud del 100%; por lo cual a partir de estos resultados solo se trabajó con la cepa LBIC-004.

La secuencia obtenida del gen *16S rRNA* de la cepa LBIC-004 fue analizada mediante Blastn y mostró un 99% de identidad con cepas del grupo de *B. cereus*; entre ellas *B. cereus*, *B. thuringiensis*, *B. mycooides*, *B. weihenstephanensis* y *B. pseudomycooides*. Sin embargo, como se mencionó en los resultados, esta secuencia tuvo un Max score con cepas de *B. cereus*. Cabe mencionar que el gen del *16S rRNA* no es una herramienta exacta para

determinar la especie dentro del grupo *cereus*. Chen y Tsen (2002) evaluaron la posibilidad de discriminación de *Bacillus cereus* y *Bacillus thuringiensis* mediante cebadores de PCR dirigidos a la región V1 del gen *16S rRNA* y *gyrB*, sin embargo, los resultados mostraron una falla en la discriminación. Del mismo modo, aunque Giffel y cols (1997) han sugerido el uso de sondas basadas en *16S rRNA* para la diferenciación de *B. cereus* y *B. thuringiensis*, dos cepas de *B. thuringiensis* fueron reconocidas como *B. cereus* y una cepa de *B. cereus* fue identificada como *B. thuringiensis*. Por otro lado, Yamada y cols. (1999) han demostrado que ocho de 10 aislados determinados como cepas de *B. thuringiensis* por la formación de inclusiones cristalinas generaron resultados de PCR positivos cuando se usaron iniciadores específicos de *B. cereus*. Por lo tanto, es necesario utilizar otros genes o herramientas para la identificación correcta de las especies dentro del grupo de *Bacillus cereus*.

Mediante análisis de Blastx de las secuencias obtenidas de la amplificación del gen *gyrB* traducida a aminoácidos se confirmó el mayor porcentaje de similitud para las seis cepas con *B. cereus* con un 100% de identidad. Sin embargo, a pesar de que se sabe que el gen *gyrB* tiene mayor resolución que el gen *16S rRNA*, por sí solo no es una herramienta que sea capaz de diferenciar entre cepas de *B. thuringiensis* y *Bacillus cereus*. Usando iniciadores basados en el gen *gyrB* diseñados para la detección de *B. cereus*, Yamada y cols (1999) obtuvieron resultados de falsos positivos para cepas de *B. thuringiensis kurstaki* y *B. thuringiensis aizawai*. Ellos encontraron que además de las cepas de *B. thuringiensis kurstaki* BT4 y BT9, una de las dos cepas de *B. thuringiensis* berliner cepa BT5, también se identificó como *B. cereus*. Aunque una cepa de *B. thuringiensis* berliner, cepa IAM 12077, ha sido identificada por Yamada como *B. thuringiensis*. Los resultados que se muestran indican que *B. thuringiensis* Berliner puede ser reconocido como cepas de *B. cereus* o *B. thuringiensis*.

La relación taxonómica entre *B. cereus* y *B. thuringiensis* todavía no está del todo clara. A pesar de que *B. cereus* puede recibir plásmidos y por lo tanto

convertirse en cepas *cry*⁺, de la misma manera *B. thuringiensis* puede perder plásmidos y convertirse en cepas *cry*⁻; la taxonomía actual los coloca en especies separadas. Por otro lado, algunos trabajos recientes propusieron que *B. cereus* y *B. thuringiensis* deben considerarse como una especie sobre la base de evidencia genética (Helgason *et al.*, 2000). Por lo tanto, todavía hay mucho debate sobre la discriminación entre cepas de *B. cereus* y *B. thuringiensis*. Los resultados del PCR y la evidencia de las secuencias mostradas por Chen y Tsen (2002) indican que, a excepción del gen *cry* o el ensayo de proteínas cristalinas, es difícil desarrollar un método molecular confiable para la diferenciación de cepas de *B. cereus* y *B. thuringiensis*.

Por último, como se mencionó en el apartado de resultados, el análisis de la secuencia de nucleótidos de la flagelina solo se realizó para la cepa LBIC-004, la cual tiene un 98% de identidad con *B. thuringiensis* serovar *entomocidus* y un 97% con *B. cereus*. Por otro lado, la secuencia aminoácidica de la flagelina mostró un 96% para *B. cereus* y *B. thuringiensis* serovar *entomocidus*, sin embargo, Reyes-Ramírez e Ibarra (2008) reportaron el patrón de plásmidos del serovar *entomocidus*, el cual carece de plásmidos pequeños y solo cuenta con un mega plásmido. Si comparamos nuestro perfil de plásmidos obtenido de la cepa LBIC-004 y lo comparamos con el patrón de *B. thuringiensis* serovar *entomocidus*, podemos determinar claramente que nuestra cepa es totalmente diferente y no pertenece a *B. thuringiensis* serovar *entomocidus*. Por lo tanto, el patrón de plásmidos junto con el análisis de los genes *16s rRNA*, *gyrB* y *hag* nos indica que efectivamente la cepa LBIC-004 pertenece a la especie *B. cereus*.

Por otra parte, la cepa LBIC-004 no es capaz de formar inclusiones parasporales durante la fase de esporulación; sin embargo, se logró detectar un gen *cry* totalmente nuevo de tres dominios el cual codifica para una proteína de 74.9 kDa. Teniendo en cuenta estos resultados y comparando con los resultados obtenidos por Ruiu y cols. (2015) para la clasificación de su cepa, los cuales son

muy similares a los obtenidos en nuestro trabajo, podemos considerar a esta cepa LBIC-004 como un nuevo aislado de *B. cereus*.

En cuanto a la distribución de genes *cry* en cepas diferentes a *Bt*, experimentalmente no encontramos ninguna distribución, ya que solo se logró identificar un gen *cry*. Sin embargo, al igual que los reportes sobre genes *cry* en otras especies, curiosamente nuestro gen es un gen *cry* de primer rango de tres dominios. Si realizamos una comparativa de los resultados obtenidos en este trabajo y en el reportado por Noguera e Ibarra (2010), donde evaluaron con los tres sets de iniciadores un total 27 cepas de *B. thuringiensis*, las cuales, sí mostraban cristales atípicos durante su fase de esporulación, ellos lograron obtener un total de 66 secuencias *cry* de las cuales solo siete genes *cry* fueron secuenciados completamente y solo tres de ellos fueron totalmente nuevos. Por otro lado, en nuestro trabajo solo identificamos un gen *cry* totalmente nuevo en cepas que no eran capaces de formar cristal durante la fase de esporulación y que no pertenecen a la especie de *B. thuringiensis*. Creemos que el resultado obtenido de esta búsqueda es bastante interesante en cuanto a la identificación de genes *cry* totalmente nuevos, ya que se partió de cepas que no nos daban indicio de que tuvieran al menos un gen *cry*. Además, podemos observar que hay una tendencia a encontrar genes *cry* totalmente nuevos en otras especies, ya que, en nuestro trabajo, al igual que en los reportes anteriores, los genes *cry* encontrados en otras bacterias siempre son de primer rango (nuevos).

En cuanto a la caracterización del gen *cry-like*, éste tiene un tamaño de 1953 pb y codifica para una proteína que tiene un 41% de identidad con la proteína Cry8Ca1 y un peso molecular de 74.9 kDa. Una de las principales características de esta proteína es que carece del extremo C-terminal, es decir, es una proteína Cry truncada. Además, ésta presenta una estructura tridimensional predicha de tres dominios muy semejante a la estructura de la proteína Cry3Aa1. A pesar de que la proteína Cry-like presenta una estructura similar a las proteínas Cry de tres dominios, ésta no tiene la capacidad de formar cristales durante la fase

de esporulación, por lo cual, podría decirse que requiere de un extremo carboxilo terminal o de un sistema de expresión más fuerte como fue el caso de la proteína Cry3A (Park *et al.*, 1998). Una característica muy peculiar de la proteína Cry3A es que a pesar de carecer del extremo C-terminal, sí tiene la capacidad de formar una inclusión proteica, aunque en muy baja cantidad. Park y *col.s.* (1998) lograron una expresión de 12.7 veces mayor de la proteína Cry3A comparado con la cepa silvestre DSM 2803; esto debido a que usaron promotores tipo *cyt1Aa* dependientes de la esporulación para dirigir la expresión de *cry3Aa* cuando se incluye la secuencia STAB-SD en la construcción.

Como se mencionó anteriormente, la proteína Cry-like encontrada no fue capaz de cristalizar por sí misma en la cepa LBIC-004, por lo tanto, como se mencionó en Materiales y Métodos, se transformó en la cepa 4Q7 de *Bt* con el vector pSTAB, el cual cuenta con la región estabilizadora y el promotor utilizado por Park y *col.s.* (1998), sin embargo, no fue capaz de formar inclusiones proteicas. Por lo tanto, se considera que para poder expresarse y cristalizar esta proteína requiere utilizar otras proteínas accesorias.

Por otra parte, a pesar de que la cepa LBIC-004 no fue capaz de cristalizar la proteína, si mostró una banda en un gel de SDS-PAGE del tamaño esperado de 74.9 kDa, lo que indicó que la proteína si es capaz de expresarse, aunque quizá en baja cantidad, y por consiguiente no es capaz de cristalizar. Por otra parte, ya que ésta no fue capaz de cristalizar bajo la expresión de un promotor fuerte y una región estabilizadora, quizá requiera de la región C-terminal para su cristalización. Además, existen reportes de proteínas Cry que requieren de una región C-terminal para su cristalización (Barboza-Corona *et al.*, 2012 y Sun *et al.*, 2013), inclusive el grupo de Peng y *col.s.* (2015) encontró que la expresión y cristalización de la proteína Cry65Aa requiere de dos regiones C-terminales para su su estabilidad, expresión y cristalización.

Por otra parte, en la cepa 4Q7-cry-like no se observó la banda del tamaño esperado; sin embargo, se identificó una banda sobre-expresada de 25 kDa que corresponde a la secuencia parcial de aminoácidos de la proteína Cry-like. Este resultado podría dar un indicio de que a pesar de que se utilizó el vector reportado por Park y cols. (1998), sí está sobre-expresando la proteína, pero ésta requiere de otros elementos que probablemente solo se encuentran en la cepa nativa (*B. cereus*), los cuales están funcionando para mantener su integridad. Si bien, existen muchos reportes donde se explica que existen otros elementos que han favorecido la estabilidad, la síntesis y la cristalización de las proteínas Cry, además de evitar su degradación, entre estos elementos podemos encontrar dominios de cristalización separados de los marcos de lectura abiertos (ORFs) como es el caso de algunas proteínas como la Cry10A, Cry19A, Cry24B y Cry65Aa1 (Hernández-Soto *et al.*, 2009; Braboza-Corona *et al.*, 2012; Ohgushi *et al.*, 2005; Peng *et al.*, 2015); además, existen factores de cristalización conocidos como P19 y P20 (Wu y Federici, 1993; Manasherob *et al.*, 2001), toxinas con factores de cristalización putativos como la Cry6A, Cry9Ca y Cry18Aa (Yu *et al.*, 2008; Lambert *et al.*, 1996; Zhang *et al.*, 1998) y por último aquellas proteínas como la Cry3A que no tiene factores de cristalización conocidos (Adalat *et al.*, 2017). Como ya se mencionó, la cepa 4Q7 transformada es capaz de expresar la proteína Cry-like, sin embargo, al parecer está siendo degradada a una proteína de menor peso molecular (25 kDa), por lo cual se hipotetiza que podría requerir de algún elemento semejante a los anteriormente mencionados, como la proteína P20 o el extremo C-terminal o cualquier otro elemento de la cepa nativa que contribuya a su estabilidad y su posible cristalización.

Por otro lado, este reporte describe una técnica alternativa para la localización de proteínas Cry, pero a través del desarrollo de un modelo para la identificación *in silico* de genes *cry* que contengan su región carboxilo terminal. A pesar de que este modelo deja fuera todas aquellas proteínas que carecen de su extremo C-terminal, este nos asegura que al menos aquellas proteínas que son detectadas mediante el modelo, todas son proteínas Cry de tres dominios. Si bien

existe un trabajo similar desarrollado por el grupo de Ye y *cols.* (2012), en el cual desarrollaron un modelo similar para la identificación de proteínas Cry, este tiene la desventaja que obtiene muchos falsos positivos y esto se debe a que en el desarrollo del modelo se incluye el extremo amino-terminal, el cual es altamente variable, a pesar de que conservan todas las proteínas una estructura de tres dominios. Por lo tanto, nuestro modelo tiene la ventaja de no detectar falsos positivos, además de que podría detectar proteínas totalmente nuevas que contengan su extremo carboxilo terminal, ya que no se sesga a las proteínas existentes.

El perfil HMM fue validado utilizando principalmente el genoma del mega plásmido pBtoxis de *Bacillus thuringiensis israelensis* el cual contiene solo dos proteínas Cry con su carboxilo terminal, la Cry4Aa y Cry4Ba y dos proteínas que carecen de éste como son la Cry10 y Cry11 (Berry *et al.*, 2002). El modelo detectó solamente las proteínas Cry4 como era de esperarse y no la Cry10 y Cry11, lo cual indicó que el modelo es altamente específico y que solo detectará aquellas proteínas que contengan su extremo carboxilo terminal. Sin embargo, también detectó tres anotaciones identificadas y reportadas por Berry y *cols.* (2002) como pBt025, pBt026 y pBt048. Las dos primarias (pBt025 y pBt026) corresponden a un fragmento similar a la Cry28Aa, esto podría indicar que el modelo es capaz de identificar si en el genoma analizado pudiera haber fragmentos de genes *cry*, lo cual ya está bien documentado que algunas cepas pueden tener en su genoma algunos genes *cry* incompletos (Ohgushi *et al.*, 2005; Hernández-Soto *et al.*, 2009; Noguera e Ibarra 2010; Barboza-Corona *et al.*, 2012; Sun *et al.*, 2013). En cuanto a la anotación de pBt048, éste es similar a la parte carboxilo de la proteína Cry4Ba. Además, como control negativo se utilizó el genoma de la cepa acristalífera de *Bacillus thuringiensis israelensis* cepa 4Q7, la cual se encuentra curada de plásmidos. En este caso el modelo no detectó ninguna coincidencia, lo cual era de esperarse ya que esta cepa no contiene ningún gen *cry*.

Con el perfil HMM se logró analizar un total de 857 genomas dentro del género *Bacillus* en los cuales se lograron identificar 174 proteínas Cry distribuidas en cuatro especies. Como era de esperarse, la mayoría de las proteínas Cry fueron identificadas en *Bacillus thuringiensis*, con un total de 132 genes encontrados en diferentes serovares de esta especie, 36 secuencias fueron identificadas en *Bacillus cereus*, un gen *cry14Aa* en *Bacillus subtilis* y 5 secuencias encontradas en cepas de *Bacillus* spp. Al analizar los resultados obtenidos mediante nuestro modelo, se observó que al igual que el trabajo de Bravo y cols. (1998) los genes *cry1* fueron los más frecuentemente encontrados con 15 subgrupos principales, el segundo gen más abundante es el gen *cry4* con 4 subgrupos, seguido del gen *cry7*. Sin embargo, cabe resaltar que como era de esperarse, hubo una tendencia de ciertos grupos de genes *cry*, ya que solo podrían ser identificados aquellos que tengan extremo carboxilo terminal.

Por otra parte, existe el inconveniente de que las bacterias cuyos genomas están en GenBank no siempre están bien clasificadas (Liu *et al.*, 2015). Por lo tanto, para corroborar que las 42 secuencias encontradas en las cepas de *B. cereus*, *B. subtilis* y *B. spp.* no pertenecieran a *B. thuringiensis*, fue que se realizó un análisis de MLST para estos genomas. El análisis de MLST es una herramienta excelente para estudiar la estructura de la población y la epidemiología del grupo *B. cereus* (Helgason *et al.*, 2004). El análisis de MLST se realizó no solo para las bacterias identificadas como *B. cereus*, sino que se realizó para los genomas que contenían al menos un gen *cry*. Con los resultados de este análisis se obtuvieron resultados muy similares a los reportados por Liu y cols. (2015). En dicho trabajo se analizaron un total de 224 genomas mediante diversas técnicas y análisis filogenéticos, entre estos análisis incluyeron el MLST con el cual encontraron resultados muy similares a los obtenidos mediante el enfoque de Genome BLAST Distance Phylogeny (GBDP) basado en la secuencia del genoma completo, con lo cual determinaron que el análisis del MLST es una herramienta muy poderosa para la clasificación de *Bacillus* dentro del grupo *cereus*. Sin embargo, uno de los resultados más importantes de Liu y cols. (2015), es que proponen que dos

especies debieran fusionarse ya que, de acuerdo con los análisis filogenéticos desarrollados en su trabajo, la especie *B. weihenstephanensis* y la *B. mycooides* pertenecen a la misma especie. Posteriormente, en el 2018, nuevamente Liu y cols. publican un análisis basado en la reclasificación del genoma de *Bacillus weihenstephanensis* como un sinónimo heterotípico de *B. mycooides*. En su trabajo, Liu (2018) concluye que no hay justificación para mantener un estado de especie diferente para *B. weihenstephanensis* y *B. mycooides*; además, debido a que el nombre *B. mycooides* tiene prioridad de publicación sobre *B. weihenstephanensis* y en base a la evidencia presentada en su estudio, proponen a *B. weihenstephanensis* como un sinónimo heterotípico de *B. mycooides* y la corrección de las identificaciones erróneas de especies de algunas bacterias relacionadas. De hecho, como se muestra en la Figura 14 de nuestro trabajo, efectivamente mediante el análisis MSLT y la reconstrucción filogenética, las dos especies se encuentran formando un solo clado.

Por otra parte, en el mismo trabajo reportado de Liu y cols. (2015) argumentan que los plásmidos portadores de genes tóxicos, tales como pXO en *B. anthracis* y los plásmidos del gen *cry* en *B. thuringiensis*, no pueden servir como firmas de ninguna de las especies. De hecho, usando dDDH (Digital DNA:DNA hybridization), algunas cepas previamente identificadas como *B. cereus* o *B. thuringiensis* se identificaron como *B. anthracis*. Además, proponen que se debe prestar mucha más atención a estas bacterias para que vuelvan a evaluar su bioseguridad, especialmente las bacterias que portan genes *cry* cuando se usan como bioplaguicidas. En resumen, Liu y cols. determinaron a partir de su estudio a gran escala del controvertido grupo *B. cereus*, que dos especies fueron reconocidas como sinónimos heterotípicos, mientras que 19-20 nuevas especies fueron reveladas y esperan una mayor caracterización. Por otra parte, varias cepas mal identificadas como *B. cereus* podrían ser reclasificadas como *Bacillus mycooides*. Además, que el grupo representado por *B. cereus* contenía 34 cepas definidas previamente como *B. cereus*, *B. thuringiensis* y *B. sp.* En cuanto al grupo representado por *B. thuringiensis* contenía catorce cepas que fueron previamente

descritas como *B. cereus* y *B. thuringiensis*. Resultados muy similares fueron encontrados en este trabajo mediante el análisis MLST y la reconstrucción filogenética, por ejemplo, la cepa JRS10 identificada en el GenBank como *B. subtilis*, de acuerdo con el análisis aquí realizado, entra dentro del grupo de *B. cereus*, por el contrario, las dos cepas registradas como *Bacillus* spp. se encuentran dentro del grupo de *B. thuringiensis*. Sorprendentemente 5 genomas previamente clasificados como *B. thuringiensis* y *B. cereus* se agruparon con la cepa tipo de *B. anthracis*. Dentro del grupo de *B. mycooides* solo se agrupó una cepa mal clasificada como *B. cereus*; de la misma manera una cepa previamente clasificada como *B. thuringiensis* se agrupó con la cepa tipo de *B. toyonensis*. Por último, muchas cepas clasificadas como *B. cereus* se agruparon con las cepas de *B. thuringiensis* y viceversa. Para finalizar, de acuerdo con los resultados encontrados mediante el análisis de MLST de los genomas seleccionados por contener al menos un gen *cry*, se logró corroborar que la distribución de los genes *cry* no es frecuente en otras especies de *Bacillus* y que probablemente los reportes que existen podrían ser de casos fortuitos. Sin embargo, de acuerdo con la posible reclasificación del grupo de *Bacillus cereus*, los genes *cry* al menos podrían estar distribuidos dentro del mismo grupo. Finalmente, a pesar del estado de confusión que existe en la actualidad sobre la diferenciación de las especies del grupo *cereus*, nuestro trabajo indica que, si bien todas las proteínas encontradas con el modelo HMM pertenecían a especies de este grupo, su distribución en otras especies diferentes a *Bacillus thuringiensis* fue definitivo. Este descubrimiento hace una contribución importante a la amplia distribución que tienen las proteínas Cry, al menos en el grupo *cereus*. La implicación de este descubrimiento abre una amplia gama de posibilidades de estudio sobre el papel de estas proteínas Cry en la naturaleza, sobre todo porque, por un lado, son más ubíquas de lo que se pensaba hasta ahora, lo cual está en concordancia con su gran diversidad, y por otro, queda dudosa la hipótesis de que las proteínas Cry sólo presentan actividad insecticida. Es necesario dedicar más trabajo al papel (nicho) que tienen estas proteínas en la naturaleza, así como el de *B. thuringiensis*.

CONCLUSIÓN

Hasta la fecha hay pocos reportes sobre la existencia de genes *cry* de tres dominios en bacterias diferentes a *Bacillus thuringiensis*. Sin embargo, los genes que han sido encontrados y reportados en bacterias que no son *Bt* tienen la característica de ser totalmente nuevos, es decir, son genes *cry* de primer rango. En el presente estudio se hizo una búsqueda exhaustiva de genes *cry* en un total de 223 cepas de *Bacillus* diferentes a *Bt*. De acuerdo con los reportes existentes y a los resultados obtenidos experimentalmente en este trabajo se puede decir que la distribución de los genes *cry* en bacterias diferentes a *Bacillus thuringiensis* no es tan común, sin embargo, como ya se menciona, se hizo una búsqueda más amplia con la ayuda de un perfil HMM y de acuerdo con los resultados obtenidos de la búsqueda de proteínas Cry en 857 genomas de *Bacillus* con 89 especies diferentes, se puede concluir que las proteínas Cry se encuentran ampliamente distribuidas dentro del grupo *cereus*.

Con la caracterización de la cepa LBIC-004 y del gen *cry-like* se logró identificar un gen *cry* totalmente nuevo, el cual pertenece a la familia de genes *cry* de tres dominios, con un 41% de identidad con la proteína Cry8Ca. En cuanto a la cepa LBIC-004, de acuerdo con los resultados obtenidos del análisis del patrón de plásmidos y el análisis de los genes *16S rRNA*, *gyrB* y *hag* se determinó que es una cepa de *Bacillus cereus*.

Por otra parte, a pesar de que los genes *cry* presentan la misma estructura terciaria en su núcleo tóxico, algunas presentan diferencias en su estabilidad y formación del cristal. Una de las características de la proteína Cry-like es que no tiene la capacidad de formar cristales y al parecer es degradada después de ser clonada en pSTAB y transformada en cepas acristalíferas de *Bt*, por lo cual llegamos a la conclusión que podría requerir de un C-terminal o de alguna de las proteínas accesorias para su estabilidad y cristalización. Por otro lado, debido a que la proteína Cry-like es totalmente nueva, no se tienen indicios sobre qué orden

de insectos podrían ser susceptibles, sin embargo, aunque no presentó toxicidad hacia los insectos probados, no indica que no pueda tener algún tipo de toxicidad hacia otros insectos.

Además, se logró desarrollar y validar un perfil HMM para la identificación de genes *cry* completos. Usando este modelo, se analizaron 857 genomas ya reportados de *Bacillus* incluyendo los genomas de *Bt*, en los cuales se lograron identificar 174 proteínas Cry; las cuales se distribuyen en 12 grupos principales. Cabe resaltar que 42 proteínas Cry de las 174 que se identificaron se encontraban en cepas diferentes a *B. thuringiensis*, lo cual es muy interesante, ya que hasta la fecha no existe ningún reporte de una búsqueda exhaustiva de proteínas Cry en otras especies y con resultados similares, sin embargo, como ya se mencionó, muchas veces los genomas no están bien clasificados en la base de datos del GenBank, por lo cual se hizo uso del análisis de MLST, con este análisis se demostró que todos los genomas que contienen al menos una secuencia de un gen *cry* pertenece a cualquier especie dentro del grupo *cereus*, pero no hubo ninguna coincidencia con otra especie de *Bacillus*. Por otra parte, uno de los resultados más importantes obtenido mediante este análisis de MLST fue que el número de genomas y por consiguiente el número de proteínas Cry presentes en cepas diferentes a *B. thuringiensis*, aumentó drásticamente de 42 a 119 proteínas Cry presentes principalmente en *B. cereus*, *B. anthracis*, *B. mycoides* y *B. toyonensis*

Finalmente, este es el primer estudio donde se hace una búsqueda exhaustiva de genes *cry* y/o proteínas Cry en bacterias diferentes a *B. thuringiensis*, con la finalidad de ver la distribución de estos genes dentro del mismo género de bacterias, sin embargo, de acuerdo con los resultados obtenidos, se demostró que los genes *cry* están ampliamente distribuidos en las especies de *Bacillus* que pertenecen al grupo *cereus*.

PERSPECTIVAS

- Identificar si el gen *cry-like* se encuentra en un plásmido o en el cromosoma.
- Agregar un C-terminal a la proteína Cry-like y determinar si es esencial para su cristalización.
- Co-expresar la proteína Cry-like con chaperonas que permitan su cristalización.
- Realizar bioensayos con la proteína cristalizada con más especies de insectos.
- Ampliar la búsqueda de proteínas cry en baciláceas y otros géneros de bacterias.

REFERENCIAS

- Adalat R., Saleem F., Crickmore N., Naz S. y Shakoory A. R. 2017. In Vivo Crystallization of Three-Domain Cry Toxins. *Toxins (Basel)*. 9(3).
- Argôlo-Filho R. C. y Loguercio L. L. 2014. *Bacillus thuringiensis* Is an Environmental Pathogen and Host-Specificity Has Developed as an Adaptation to Human-Generated Ecological Niches. *Insects*, 5(1), 62–91.
- Bagos P. G., Liakopoulos T. D, Spyropoulos I. C y Hamodrakas S. J. 2004. A Hidden Markov Model method, capable of predicting and discriminating beta-barrel outer membrane proteins. *BMC Bioinformatics*. 15; 5:29.
- Barboza-Corona J. E., López-Meza J. E. e Ibarra J.E. 1998. Cloning and expression of the cry1Ea4 gene of *Bacillus thuringiensis* and the comparative toxicity of its gene product. *World J Microb Biot.*14, 437–41.
- Barboza-Corona J. E., Park H. W., Bideshi D. K. y Federici B. A. 2012. The 60-kilodalton protein encoded by orf2 in the cry19A operon of *Bacillus thuringiensis* subsp. jegasan functions like a C-terminal crystallization domain. *Appl Environ Microbiol.* 78 (6): 2005-12.
- Barloy F., Delécluse A., Nicolas L. y Lecadet M. M. 1996. Cloning and expression of the first anaerobic toxin gene from *Clostridium bifermentans* subsp. malaysia, encoding a new mosquitocidal protein with homologies to *Bacillus thuringiensis* delta-endotoxins. *Journal of Bacteriology*, 178(11), 3099–3105.
- Barloy F., Lecadet M. M. y Delécluse A. 1998. Cloning and sequencing of three new putative toxin genes from *Clostridium bifermentans* CH18. *Gene*, 211(2), 293–299.

- Bechtel D.B. y Bulla L. A. 1982. Ultrastructural analysis of membrane development during *Bacillus thuringiensis* sporulation. *J Ultrastruct Res.* 79:121-32.
- Berliner E. Åœber die schlafsucht der mehlmottenraupe. *Z Gesamte Getreidewes.* 1911 ; 3:63. (Ita).
- Berry C., O'Neil S., Ben-dov E., Jones A. F., Murphy L., Quail M. A, Holden M.T., Harris D., Zaritsky A. y Parkhill J. 2002. Complete Sequence and Organization of pBtoxis, the Toxin-Coding Plasmid of *Bacillus thuringiensis* subsp. israelensis. *Applied and Environmental Microbiology*, 68(10), 5082–5095.
- Bravo A., Sarabia S., Lopez L., Ontiveros H., Abarca C., Ortiz A, Ortiz M., Lina L., Villalobos F. J., Peña G., Nuñez-Valdez M. E., Soberón M. y Quintero R. 1998. Characterization of cry Genes in a Mexican *Bacillus thuringiensis* Strain Collection. *Appl Environ Microbiol.* 64 (12): 4965-72.
- Bravo A., Gill S. S. y Soberon M. 2007. Mode of action of *Bacillus thuringiensis* cry and cyt toxins and their Potential for Insect Control. *Toxicon*, 49(4), 423–435.
- Bravo A., Gómez I., Porta H., García-Gómez B. I., Rodriguez-Almazan C., Pardo L. y Soberón M. 2013. Evolution of *Bacillus thuringiensis* Cry toxins insecticidal activity. *Microbial Biotechnology*, 6(1), 17–26.
- Bulla L. A., Bechtel D. B., Kramer K. J., Shethna Y. I., Aronson A. I. y Fitz-James P.C. Ultrastructure, physiology and biochemistry of *Bacillus thuringiensis*. *Crit Rev. Microbiol*, 8(2): 147-204.
- Chen M. L. y Tsen H. Y. 2002. Discrimination of *Bacillus cereus* and *Bacillus thuringiensis* with 16S rRNA and *gyrB* gene based PCR primers and sequencing of their annealing sites. *Journal of Appl. Microbio.* 92, 912–919.

- Crickmore, N., Baum, J., Bravo, A., Lereclus, D., Narva, K., Sampson, K., Schnepf, E., Sun, M. y Zeigler, D.R. "*Bacillus thuringiensis* toxin nomenclature" (2018) <http://www.btnomenclature.info/>
- Crickmore N. 2000. The diversity of *Bacillus thuringiensis* δ -endotoxins. In: Charles JF., Delécluse A., Roux C.NL. (eds) Entomopathogenic Bacteria: from Laboratory to Field Application. Springer, Dordrecht
- Crickmore N., Zeigler D. R., Feitelson J., Schnepf E., Van Rie J., Lereclus D., Baum J. y Dean D. H. 1998. Revision of the nomenclature for the *Bacillus thuringiensis* pesticidal crystal proteins. *Microbiology and Molecular Biology Reviews*: 62(3), 807–813.
- De Maagd R. A., Bravo A. y Crickmore N. 2001. How *Bacillus thuringiensis* has evolved specific toxins to colonize the insect world. *Trends in Genetics*, 17(4), 193–199.
- Deist B. R., Rausch M. A., Fernandez-Luna M. T, Adang M. J. y Bonning B. C. 2014. Bt toxin modification for enhanced efficacy. *Toxins*, 6(10), 3005–3027.
- Durbin R., Eddy S. R., Krogh A. y Mitchison G. J. 1998. *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*. Cambridge University Press, Cambridge UK.
- Eddy S. R. 1998. Profile hidden Markov models. *Bioinformatics*, 14:755–763.
- Evdokimov A. G., Moshiri F., Sturman E. J., Rydel T. J., Zheng M., Seale J. W. y Franklin S. 2014. Structure of the full-length insecticidal protein Cry1Ac reveals intriguing details of toxin packaging into in vivo formed crystals. *Protein Science*, 23(11), 1491–1497.

- Fang Y., Li Z., Liu J., Shu C., Wang X., Zhang X., Yu X., Zhao D., Liu G., Hu S., Zhang J., Al-Mssallem I. y Yu, J. 2011. A pangenomic study of *Bacillus thuringiensis*. *J. Genet. Genomics*. 38, 567–576.
- Farrar M. 2007. Striped Smith-Waterman speeds database searches six times over other SIMD implementations. *Bioinformatics*. 23(2): 156-61.
- Fu L., Niu B., Zhu Z., Wu S. y Li W. 2012. CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics*, 28(23), 3150–3152.
- Giffel M. C., Beumer R. R., Klijn N., Wagendorp A. y Rombouts F. M. 1997. Discrimination between *Bacillus cereus* and *Bacillus thuringiensis* using specific DNA probes based on variable regions of 16S rRNA. *FEMS Microbiology Letters* 146, 47–51.
- Ge A. Z., Shivarova N. I. y Dean D. H. 1989. Location of the *Bombyx mori* specificity domain on a *Bacillus thuringiensis* delta endotoxin protein. *Proc. Natl. Acad. Sci. USA*, 86, 4037–4041.
- Gómez-Lunar Z., Hernández-González I., Rodríguez-Torres M. D., Souza V. y Olmedo-álvarez G. 2016. Microevolution analysis of *Bacillus coahuilensis* unveils differences in phosphorus acquisition strategies and their regulation. *Frontiers in Microbiology*, 7, 1–16.
- Gong Y. N., Chen G.W. y Shih S. R. 2012. Characterization of subtypes of the influenza A hemagglutinin (HA) gene using profile hidden Markov models. *J Microbiol Immunol Infect*. 45(6):404-10.
- Haeyoung Jeong, Seung-Hwan Park y Soo-Keun Choi. 2014. Genome Sequence of the AcrySTALLiferous *Bacillus thuringiensis* Serovar Israelensis Strain 4Q7, Widely Used as a Recombination Host. *Genome Announc*. 2(2): e00231-14.

- Helgason E. Ø., Kstad O. A., Caugant D. A., Johansen H. A., Fouet A., Mock M., Hegna I. y Kolstø, A. B. 2000. *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis* — one species on the basis of genetic evidence. *Applied and Environmental Microbiology* 66, 2627–2630.
- Helgason E., Tourasse N. J., Meisal R., Caugant D. A. y Kolstø, A. B. 2004. Multilocus sequence typing scheme for bacteria of the *Bacillus cereus* group. *Appl Environ Microbiol.* 70 (1): 191-201.
- Hernández-Soto A., Del Rincón-Castro M. C., Espinoza A. M. e Ibarra J. E. 2009. Parasporal body formation via overexpression of the Cry10Aa toxin of *Bacillus thuringiensis* subsp. *israelensis*, and Cry10Aa-Cyt1Aa synergism. *Appl Environ Microbiol.* 75(14):4661-7.
- Hill K. K., Ticknor L. O., Okinaka R. T., Asay M., Blair H., Bliss K. A., Laker M., Pardington P. E., Richardson A. P., Tonks M., Beecher D. J., Kemp J. D., Kolstø A. B., Wong A. C., Keim P. y Jackson P. J. 2004. Fluorescent amplified fragment length polymorphism analysis of *Bacillus anthracis*, *Bacillus cereus*, y *Bacillus thuringiensis* isolates. *Appl. Environ. Microbiol.* 70:1068-1080.
- Hoffmaster A. R., Ravel J., Rasko D. A., Chapman G. D., Chute M. D., Marston C. K., De B. K., Sacchi C. T., Fitzgerald C., Mayer L. W., Maiden M. C., Priest F. G., Barker M., Jiang L., Cer R. Z., Rilstone J., Peterson S. N., Weyant R. S., Galloway D. R., Read T. D., Popovic T. y Fraser C. M. 2004. Identification of anthrax toxin genes in a *Bacillus cereus* associated with an illness resembling inhalation anthrax. *Proc Natl Acad Sci USA.* 101, 8449–8454.
- Höfte H. y Whiteley, H. R. 1989. Insecticidal crystal proteins of *Bacillus thuringiensis*. *Microbiological Reviews*, 53(2), 242–255.

- Huang Y., Niu B., Gao Y., Fu L. y Li W. 2010. CD-HIT Suite: A web server for clustering and comparing biological sequences. *Bioinformatics*, 26(5), 680–682.
- Ibarra J. E., Del Rincón M. C., Ordúz S., Noriega D., Benintende G., Monnerat R., Regis L., De Oliveira C. M., Lanz H., Rodriguez M. H., Sánchez J., Peña G. y Bravo A. 2003. Diversity of *Bacillus thuringiensis* strains from Latin America with insecticidal activity against different mosquito species. *Appl Environ Microbiol.* 69(9):5269-74.
- Ibrahim M. A., Griko N., Junker M. y Bulla L. A. 2010. A genomics and proteomics perspective. *Bioengineered Bugs*, 1(1), 31–50.
- Irriarte J. y Caballero P. 2001. Bioinsecticidas: fundamentos y aplicaciones de *Bacillus thuringiensis* en el control integrado de plagas. Cap 1, pag. 35-36.
- Jones G. W., Nielsen-Leroux C., Yang Y., Yuan Z., Dumas V. F., Gomes Monnerat R. y Berry C. 2007. A new Cry toxin with a unique two-component dependency from *Bacillus sphaericus*. *The FASEB Journal*, 21(14), 4112–4120.
- Kaelin P., Morel P. y Gadani F. 1994. Isolation of *Bacillus thuringiensis* from Stored Tobacco and *Lasioderma serricorne* (F.). *Appl Environ Microbiol.* 60(1):19-25.
- Keim P., Price L. B., Klevytska A.M., Smith K. L., Schupp J. M., Okinaka R., Jackson, P. J. y Hugh-Jones M. E. 2000. Multiple-locus variable-number tandem repeat analysis reveals genetic relationships within *Bacillus anthracis*. *Appl. Environ. Microbiol.* 182:2928-2936.

- Kelley L. A., Mezulis S., Yates C. M., Wass M. N y Sternberg M. J. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc.* 10 (6): 845-58.
- Kolstø A. B., D. Lereclus y M. Mock. 2002. Genome structure and evolution of the *Bacillus cereus* group. *Curr. Top. Microbiol. Immunol.* 264, 95-108.
- Krogh A., Brown M., Mian I. S., Sjölander K. y Hausssler D. 1994. Hidden Markov models in computational biology: Applications to protein modeling. *J. Mol. Biol.*, 235:1501–1531.
- Laemmli UK. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*; 227: 680–85.
- Lambert B., Buysse L., Decock C., Jansens S., Piens C. Saey B. Seurinck J., Van Audenhove K., Van Rie J. y Van Vliet A. A. 1996. *Bacillus thuringiensis* insecticidal crystal protein with a high activity against members of the family Noctuidae. *Appl. Environ. Microbiol.* 62, 80–86.
- Letunic I. y Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic. Acids. Res.* 44 (Web Server issue), W242–W245.
- Li W. y Godzik A. 2006. Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, 22(13), 1658–1659.
- Li W., Jaroszewski L. y Godzik A. 2001. Clustering of high homologous sequences to reduce the size of large protein database, 17(3), 282–283.
- Li W., Jaroszewski L. y Godzik A. 2002. Tolerating some redundancy significantly speeds up clustering of large protein databases. *Bioinformatics*, 18(1), 77–82.

- Liu Y., Lai Q., Göker M., Meier-Kolthoff J. P., Wang M., Sun Y., Wang L. y Shao Z. 2015. Genomic insights into the taxonomic status of the *Bacillus cereus* group. *Sci Rep.* 16; 5:14082.
- Liu Y., Lai Q. y Shao Z. 2018. Genome analysis-based reclassification of *Bacillus weihenstephanensis* as a later heterotypic synonym of *Bacillus mycoides*. *Int J Syst Evol Microbiol.* 68(1):106-112.
- Logan N. A., Turnbull P. C. y Murray P.R. 1999. Ed. Manual of clinical microbiology. American Society for Microbiology, Washington, DC. 357–369
- Loong S. N. K. y Mishra S. K. 2007. Unique folding of precursor microRNAs: quantitative evidence and implications for de novo identification. *RNA*, 13, 170–187.
- Manasherob R., Zaritsky A., Ben-Dov E., Saxena D., Barak Z. y Einav M. 2001. Effect of accessory proteins P19 and P20 on cytolytic activity of Cyt1Aa from *Bacillus thuringiensis* subsp. israelensis in *Escherichia coli*. *Curr Microbiol.* 43 (5): 355-64.
- Margulis L., Jorgensen J. Z., Dolan S., Kolchinsky R., Rainey F. A y Lo S. C. 1998. The Arthromitus stage of *Bacillus cereus*: intestinal symbionts of insects. *Proc. Natl. Acad. Sci. USA* 95:1236-1241.
- Martin P. A. W. y Travers R. S. 1989. Worldwide abundance and distribution of *Bacillus thuringiensis* isolates. *Appl. Environ. Microbiol.* 55:2437–2442.
- Melo A. L. D. A., Soccol V. T. y Soccol C. R. 2016. *Bacillus thuringiensis*: Mechanism of action, resistance, and new applications: A review. *Critical Reviews in Biotechnology*, 36(2), 317–326

- Méndez-López I., Basurto-Ríos R. e Ibarra J. E. 2003. *Bacillus thuringiensis* serovar israelensis is highly toxic to the coffee berry borer, *Hypothenemus hampei* Ferr. (Coleoptera: Scolytidae). *FEMS Microbiology Letters*, 226(1), 73–77. [https://doi.org/10.1016/S0378-1097\(03\)00557-3](https://doi.org/10.1016/S0378-1097(03)00557-3)
- Mock M. y Fouet A. 2001. Anthrax. *Annual Review of Microbiology* 55, 647–671.
- Muñoz-Medina J. E., Sánchez-Vallejo C. J., Méndez-Tenorio A., Monroy-Muñoz I. E., Angeles-Martínez J., Santos Coy-Arechavaleta A., Santacruz-Tinoco C. E., González-Ibarra J., Anguiano-Hernández Y. M., González-Bonilla C. R., Ramón-Gallegos E. y Díaz-Quiñonez J. A. 2015. In silico identification of highly conserved epitopes of influenza A H1N1, H2N2, H3N2, and H5N1 with diagnostic and vaccination potential. *BioMed Research International*. 2015:813047
- Niu B., Fu L., Sun S. y Li W. 2010. Artificial and natural duplicates in pyrosequencing reads of metagenomic data. *BMC Bioinformatics*. 13; 11, 187.
- Noguera P. A. e Ibarra J. E. 2010. Detection of new cry genes of *Bacillus thuringiensis* by use of a novel PCR primer system. *Applied and Environmental Microbiology*. 76(18) 6150-6155.
- Ohgushi A., Saitoh H., Wasano N., Uemori A. y Ohba M. 2005. Cloning and characterization of two novel genes, cry24B and s1orf2, from a mosquitocidal strain of *Bacillus thuringiensis* serovar sotto. *Curr. Microbiol.* 51,131–136.
- Okinaka R., Cloud K., Hampton O., Hoffmaster A., Hill K., Keim P., Koehler T., Lamke G., Kumano S., Manter D., Martinez Y., Ricke D., Svensson R. y Jackson P. 1999. Sequence, assembly and analysis of pXO1 and pXO2. *J. Appl. Microbiol.* 87 (2): 261-262.

- Okinaka R. T., Cloud K., Hampton O., Hoffmaster A. R., Hill K. K., Keim P., Koehler T. M., Lamke G., Kumano S., Mahillon J., Manter D., Martinez Y., Ricke D., Svensson R. y Jackson P.J. 1999. Sequence and organization of pXO1, the large *Bacillus anthracis* plasmid harboring the anthrax toxin genes. *J. Bacteriol.* 181, 6509–6515.
- Palma L., Muñoz D., Berry C., Murillo J., Caballero P. y Caballero P. 2014. *Bacillus thuringiensis* toxins: An overview of their biocidal activity. *Toxins*, 6(12), 3296–3325.
- Park H. W., Ge B., Bauer L. S. y Federici B. A. 1998. Optimization of Cry3A yields in *Bacillus thuringiensis* by use of sporulation-dependent promoters in combination with the STAB-SD mRNA sequence. *Appl Environ Microbiol.* 64(10):3932-8.
- Peng D. H., Pang C. Y., Wu H., Huang Q., Zheng J. S. y Sun M. 2015. The expression and crystallization of Cry65Aa require two C-termini, revealing a novel evolutionary strategy of *Bacillus thuringiensis* Cry proteins. *Sci Rep.* 6; 5:8291.
- Petras S. F. y Casida L. E. Survival of *Bacillus thuringiensis* spores in soil. *Appl. Environ. Microbiol.* 1985; 50: 1496–1501.
- Priest F. G., Barker M., Baillie L. W. J., Holmes E. C. y Maiden M. C. J. 2004. Population Structure and Evolution of the *Bacillus cereus* Group. *Journal of Bacteriology*, 186(23), 7959–7970.
- Reinoso-Pozo Y., Del Rincón-Castro M. C. e Ibarra J. E. 2016. Characterization of a highly toxic strain of *Bacillus thuringiensis* serovar kurstaki very similar to the HD-73 strain. *FEMS Microbiol Lett.* 363(17), 1–6.

- Restrepo-Montoya D., Becerra D., Carvajal-Patiño J. G., Mongui A., Niño L. F., Patarroyo M. E. y Patarroyo M. A. 2011. Identification of Plasmodium vivax proteins with potential role in invasion using sequence redundancy reduction and profile hidden Markov models. PLoS One. 6(10):e25189.
- Reyes-Ramírez A. e Ibarra J. E. 2008. Plasmid patterns of *Bacillus thuringiensis* type strains. Appl Environ Microb; 74:125–9.
- Rubinstein R. y Fiser A. 2008. Predicting disulfide bond connectivity in proteins by correlated mutations analysis. Bioinformatics, 24, 498–504.
- Ruiu L., Falchi G., Floris I., Marche M. G., Mura M. E. y Satta A. 2015. Pathogenicity and characterization of a novel *Bacillus cereus sensu lato* isolate toxic to the Mediterranean fruit fly *Ceratitis capitata* Wied. Journal of Invertebrate Pathology, 126, 71–77.
- Sauka D. H. y Benintende G. B. 2008. *Bacillus thuringiensis*: generalidades. Un acercamiento a su empleo en el biocontrol de insectos lepidópteros que son plagas agrícolas. Revista Argentina de Microbiología, 40(2), 124–140.
- Schnepf E., Crickmore N., Van Rie J., Lereclus D., Baum J., Feitelson J., Zeigler D.R. y Dean D. H. 1998. *Bacillus thuringiensis* and its pesticidal crystal proteins. Microbiol Mol Biol Rev. 62(3), 775–806.
- Sharp C., Bray J., Housden N. G., Maiden M. C. J. y Kleanthous C. 2017. Diversity and distribution of nuclease bacteriocins in bacterial genomes revealed using Hidden Markov Models. PLoS Comput Biol. 13(7): e1005652.
- Soufiane B. y Côté J. C. 2009. Discrimination among *Bacillus thuringiensis* H serotypes, serovars and strains based on 16S rRNA, *gyrB* and *aroE* gene sequence analyses. Antonie Van Leeuwenhoek. 95(1):33-45.

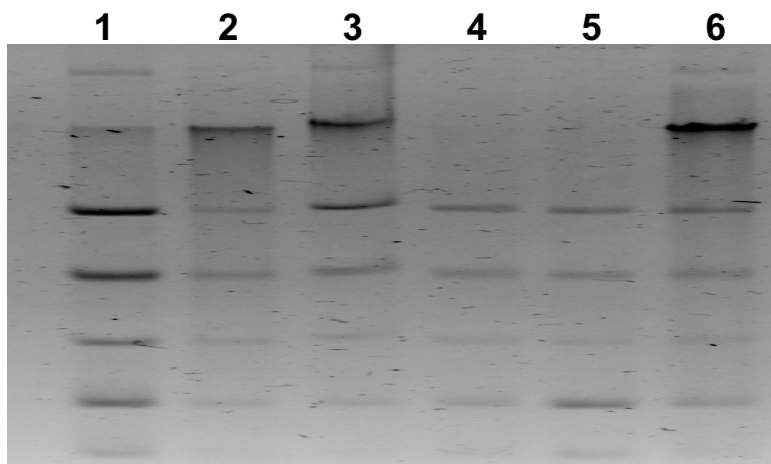
- Sun Y., Zhao Q., Xia L., Ding X., Hu Q., Federici B. A. y Park H. W. 2013. Identification and characterization of three previously undescribed crystal proteins from *Bacillus thuringiensis* subsp. *jegathesan*. *Appl Environ Microbiol.* 79 (11): 3364-70.
- Sun S., Chen J., Li W., Altintas I., Lin A., Peltier S., Stocks K., Allen E. E., Ellisman M., Grethe J. y Wooley J. 2011. Community cyberinfrastructure for Advanced Microbial Ecology Research and Analysis: the CAMERA resource. *Nucleic Acids Res.*, 39, D546–D551.
- Suzek B. E., Huang H., McGarvey F., Mazumder R y Wu CH. 2007. UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics*, 23(10), 1282–1288.
- Tabassum R. 2017. Molecular cloning and 3D model of first cytochrome P450 from CYP3A subfamily in saltwater crocodile (*Crocodylus porosus*). *Biochem Biophys Res Commun.* pii: S0006-291X(17)32057-0
- Tamura K., Peterson D., Peterson N., Stecher G., Nei M. y Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28 (10), 2731–2739.
- Turnbull P.C.B. 2002. Introduction: anthrax history, disease and ecology. *Curr Top Microbiol Immunol.* 271, 1–19.
- Walsh C. J., Guinane C. M., O' Toole P.W. y Cotter P. D. 2017. A Profile Hidden Markov Model to investigate the distribution and frequency of LanB-encoding lantibiotic modification genes in the human oral and gut microbiome. *PeerJ.* 5:e3254.

- Wei D., Jiang Q., Wei Y. y Wang S. 2012. A novel hierarchical clustering algorithm for gene sequences. *BMC Bioinformatics*, 13(1), 1.
- Weizhong L., Limin F., Beifang N., Sitao W. y John W. 2012. Ultrafast clustering algorithms for metagenomic sequence analysis. *Briefings in Bioinformatics*, 13(6), 656–668.
- Wu D. y Federici B. A. 1993. A 20-kilodalton protein preserves cell viability and promotes CytA crystal formation during sporulation in *Bacillus thuringiensis*. *J Bacteriol.* 175(16): 5276-80.
- Xu D. y Côté J. C. 2006. Sequence diversity of the *Bacillus thuringiensis* and *B. cereus sensu lato* flagellin (H antigen) protein: comparison with H serotypes diversity. *Appl Environ Microb.* 72:4653–62.
- Yamada S., Ohashi E., Agata N. y Venkateswaran K. 1999. Cloning and nucleotide sequence analysis of *gyrB* of *Bacillus cereus*, *B. thuringiensis*, *B. mycoides*, and *B. anthracis* and their application to the detection of *B. cereus* in rice. *Applied and Environmental Microbiology* 65, 1483–1490.
- Ye W., Zhu L., Liu Y., Crickmore N., Peng D., Ruan L. y Sun M. 2012. Mining new crystal protein genes from *Bacillus thuringiensis* on the basis of mixed plasmid-enriched genome sequencing and a computational pipeline. *Appl Environ Microbiol.* 78(14):4795-801.
- Yokoyama T., Tanaka M. y Hasegawa M. 2004. Novel cry gene from *Paenibacillus lentimorbus* strain Semadara inhibits ingestion and promotes insecticidal activity in *Anomala cuprea* larvae. *Journal of Invertebrate Pathology*, 85(1), 25–32.

- Yooseph S., Li W. y Sutton G. 2008. Gene identification and protein classification in microbial metagenomic sequence data via incremental clustering. *BMC Bioinformatics*. 10;9, 182.
- Yu Z., Bai P., Ye W., Zhang F., Ruan L. y Sun M. 2008. A novel negative regulatory factor for nematocidal Cry protein gene expression in *Bacillus thuringiensis*. *J. Microbiol. Biotechnol.* 18, 1033–1039
- Zhang L., Peng Y., Wu S., Sun L., Huang E., Huang T., Xu L., Wu C., Gelbic I. y Guan X. 2012. Microbial ecology and association of *Bacillus thuringiensis* in chicken feces originating from feed. *Curr. Microbiol.* 65:784–791.
- Zhang Z., Schwartz S., Wagner L. y Miller W. 2000. A greedy algorithm for aligning DNA sequences. *J Comput Biol.* 7(1-2):203-14.
- Zhang J., Schairer H. U., Schnetter W., Lereclus D. y Agaisse H. 1998. *Bacillus popilliae* cry18Aa operon is transcribed by σ E and σ K forms of RNA polymerase from a single initiation site. *Nucleic Acids Res.* 26, 1288–1293.
- Zhang J., Hodgman T. C., Krieger L., Schnetter W. y Schairer H. U. 1997. Cloning and analysis of the first cry gene from *Bacillus popilliae*. *J Bacteriol.* 179(13), 4336–4341.

ANEXO 1

Patrón de plásmidos de las cepas LBIC-R6 (1), LBIC-S9 (2), LBIC-T4 (3), LBIC-T7 (4), LBIC-Y4 (5), LBIC-004 (6).



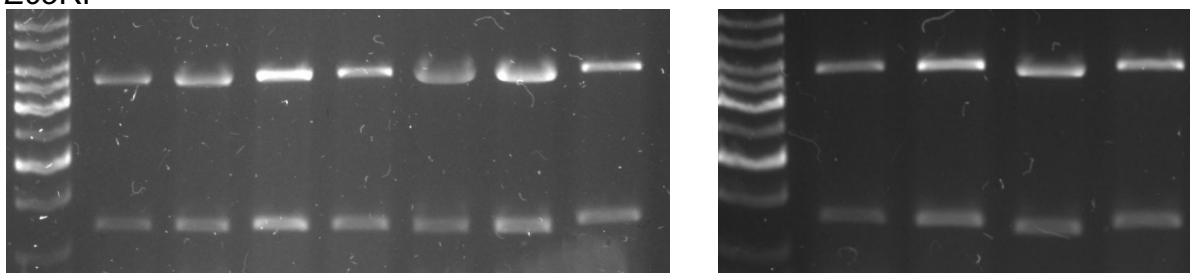
Análisis del gen *16S rRNA* y *gyrB* de las cepas LBIC-R6, LBIC-S9, LBIC-T4, LBIC-T7, LBIC-Y4, LBIC-004

Cepa	Análisis del gen <i>16S rRNA</i> ribosomal	Identidad	Rdp	Análisis del gen <i>gyrB</i> como proteína	Identidad
LBIC-R6	<i>Bacillus cereus</i> ATCC 14579	100%	Bc	DNA gyrase subunit B [<i>Bacillus cereus</i> VD014]	100%
LBIC-S9	<i>Bacillus cereus</i> strain ATCC 14579	99%	Bc	DNA gyrase subunit B [<i>Bacillus cereus</i>]	100%
LBIC-T4	<i>Bacillus cereus</i> strain CCM 2010	99%	Bc	DNA gyrase subunit B [<i>Bacillus cereus</i>]	100%
LBIC-T7	<i>Bacillus cereus</i> ATCC 14579	99%	Bc	DNA gyrase subunit B [<i>Bacillus cereus</i>]	100%
LBIC-Y4	<i>Bacillus cereus</i> strain JCM 2152	99%	Bc	DNA gyrase subunit B [<i>Bacillus cereus</i>]	100%
LBIC-004	<i>Bacillus cereus</i> ATCC 14579	99%	Bc	DNA gyrase subunit B [<i>Bacillus cereus</i>]	100%

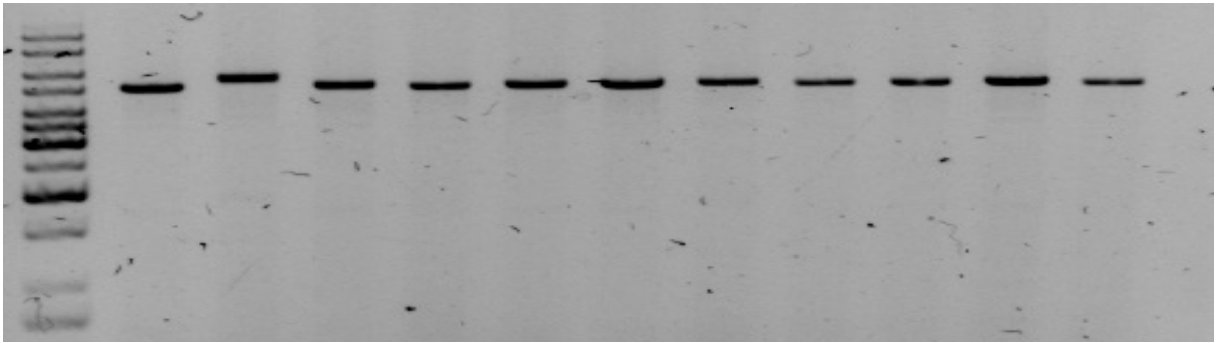
ANEXO 2

Análisis de restricción con diferentes enzimas de las clonas obtenidas del amplicon de la cepa LBIC-004

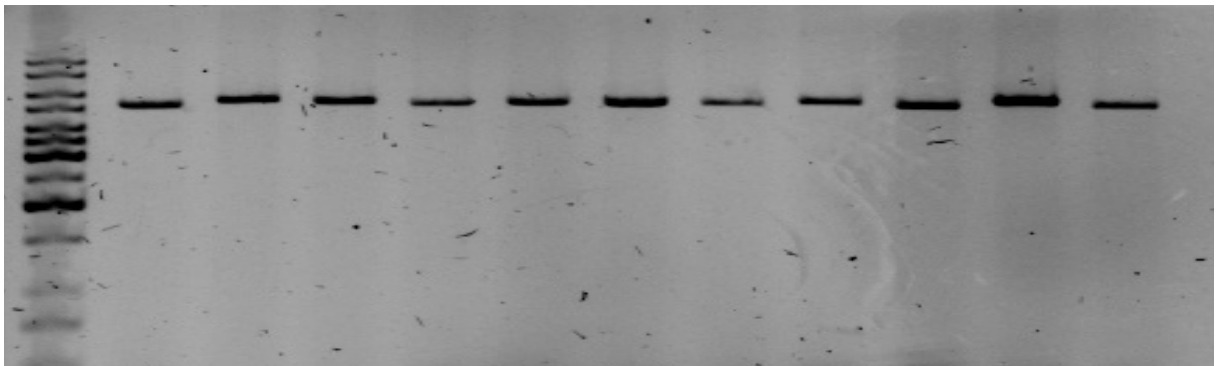
EcoRI



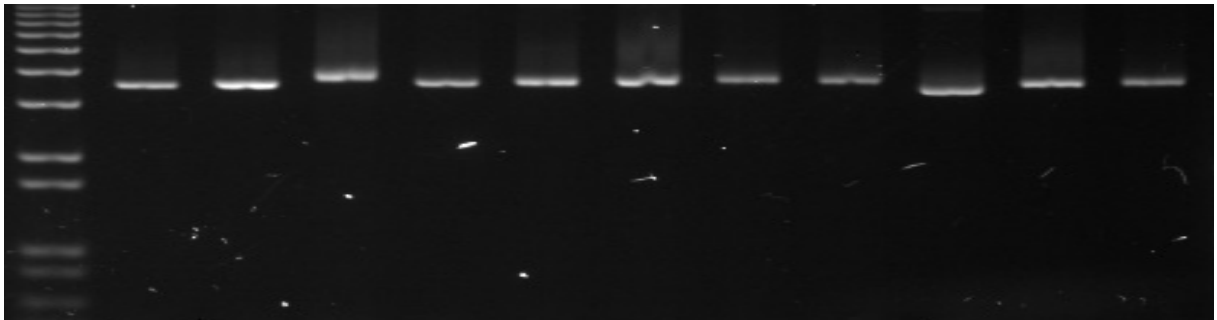
EcoRV



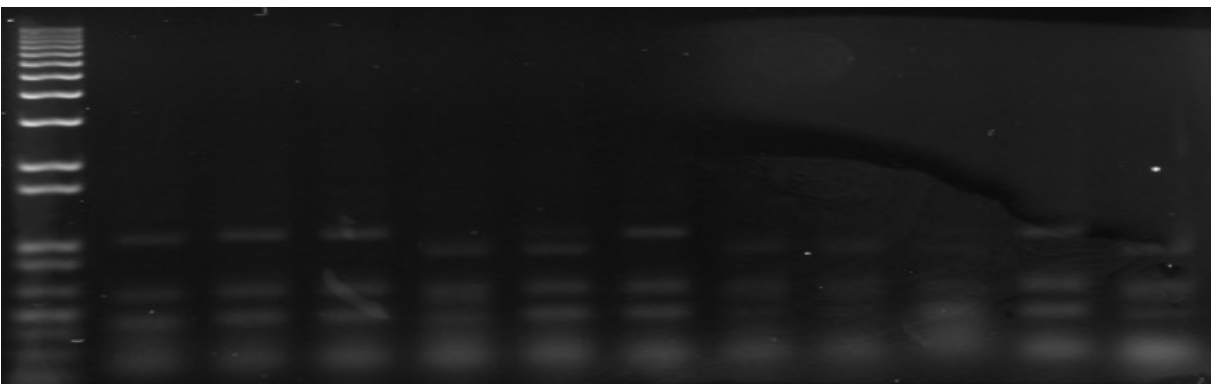
BamHI



HindIII



AluI



ANEXO 3

Secuencia del nuevo gen *cry-like* de la cepa LBIC-004

Secuencia nucleotídica 1953pb

ATGAATTCAAAGACGAAAACAACTTAATACCCTAAACAATTCATCTAATCAATCAACATCAA
TAATTCTGAACGATTTTCTTTAGCTACCGACCAAACAGTACATTACAAATCGAGAATGATGTAG
ATTATCTGAGGGTATCTGAGGGAGCTGAAAACGAGGTTTTATCGAATCATGAGGCACTTATTAG
TGCGCCAGGTCTAGTTGATATAACGGCTAATATTACTAGTACTATTCTGGGCGCTCTAGGAGTC
CCCGTTTTAGGAACTATAGTAAAATTATATAGTAAATTATTTGGTTTTTTATGGGGGGCAACACC
AGGACAAGATCCATGGAAAGAGCTAATGGATCGAGTGAAATACTCATTGAACAGAAATTAACA
GAATACGCAAGAAGTAAAGCATTAGCAGAATTGGAAGGATTGCAAAATATTATGAAATCCTATG
TGGATGCACTTGAATCATGGCAAATAACTCTCGGAATTCACAAGCAAGGTTATTAGTACAACA
GAGATTTGTTGTAGCAGATTCACACTTTAAAAGGCAATGCCTTCCTTCGCAATTAATAATTATG
AAGTATCATTGTTACCGGTATATGCACAAGCTGCAAATCTTCATCTACTTTTATTAAGAGATAGT
CAAATTTTTGGAAAAGAGTGGGGAATGCCTCAAATGAAATTGATATTTTTATAATGAACAATT
AAATCTTTAGAAAAATATTCTGACCATTGTGTAGAATGGTATCATACTGGCTTAAACCAACTAA
AAGATAAAGGATCATCTGCTAAGAACTGGGTTGATTACAATCGTTATCGCAGAGAAATGACATT
GGCAGTGCTAGATATCGTTGCGTACTTCCTAACTATGATATATATATGTACCCTATGCCGGTA
CATGCAGAACTTACGAGGGAGATATATACGGATCCGGTGGGGAGTTATTTACCCTACAAAGGT
AATTTTTATGATGTTATGTCTTGGTATGAGATGACAAAGGCATATCGCCAACCTACATTCCAAGA
ATTAGAAAATTTGATTAGGAAGCCAAGTCGGTTCACATGGATAGATAATCTTCAGATGTATACAA
GAAAAGACAGAGTAGTCAATATGAGTATTATAATTATTGGGTAGGGCTTGTGTTAAAACAATCT
CTAATTAGCGCAAGCGAGCCGAATGTAATAATTACCAGTGGTGAGATTACTAGCGAAAAAGATA
GATTTGATTTTAGTTTTATGGACCTTTACCGATATATAAAGTTATATGTAATTATATAGGGCGAT
ACGAGTCAAGTTTAGTGGGAGTAAATCAAGTTGAGTTTTATTATCTAGTAAATAATACTCCTGAG
AAACAAGAATATAAAAAAGATATTGTGGTTATGCAACAATCGCAAAAACTATTGACTCAACAGA
GGAATTGAATAATCATACGCTGAGTTTTATTGAATCCTTTGATTTATATTGGAATAATAAAAATGA
AAAAGGAGGAACAATTCCTATATTTGGATGGACACATCATAGTGTAGCTCCTAAAAATATAATTT
TTGAGGATAAGATTACACAAGTGCCAGTTATTAATCTAATGATGGGACCCAATCTGAAGTTAT
ACAAGGACCTGGATATACAGGAGGAAATTTAATAAGAGGTTCTAGATATATGGGGAATCTTACA
ATTACTCCTCCTTAGTAAATACGGAAAAAATATCGTATGAGGATTAGATATGCCGCTGATTC
AGATGGCTTACTTGAAGTGCAGATTCACAAGAACAGATCTACATATCAAATGCCCTTTAAAGCA
ACAATGAAAAAAGGGGAGGCTTTCAAATTTAATTCCTTTCAATATGTAGAAAAAGAAGTAACCTT
TGAAACTTTTCATGCTGGTATTTTTTTAAACGCGCATAATGGTCTATATTTAGATAAAATTGAGTT
TATACCTATGTAA

Secuencia aminoacídica 650aa

MNSKDENKLNLTNNSSNQSTSNNSERFSLATDQTSTLQIENDVDYLRVSEGAENEVLSNHEALISA
 PGLVDITANITSTILGALGVPVLGTIVKLYSKLFGFLWGATPGQDPWKELMDRVEILIEQKLTEYARSK
 ALAELEGLQNIMKSYVDALESWQNNRNSQARLLVQQRFFVADSHFKKAMPSSFAIKNYEVSLLPVY
 AQAANLHLLLLRDSQIFGKEWGMPQNEIDIFYNEQLNSLEKYSDHCVWEYHTGLNQLKDKGSSAK
 NWWVDYNRYRREMTLAVLDIVALLPNYDIYMYPMPVHAELTREIYTDVPVGSYLPYKGNFYDVMMSWYE
 MTKAYRQPTFQELENLIRKPSRFTWIDNLQMYTRKRQSSQYEYYNYWVGLVVKQLSASEPNVIIT
 SGEITSEKDRFDFSFYGPLPIYKVICNYIGRYESSLVGVNQVEFYLVNNTPEKQEYKKDIVVMQQS
 QKTIDSTEELNNHTLSFIESFDLYWNNKNEKGGTIPIFGWTHHSVAPKNIIFEDKITQVPVIKSNDBGTQ
 SEVIQGGYTGGNLIRGSRYMGNLTITPPSSKYGKKYRMRIRYAADSDGLLEVQIHKNRSTYQMPF
 KATMKKGAEAFKNSFQYVEKEVTLETFHAGIFLNAHNGLYLDKIEFIPM Stop

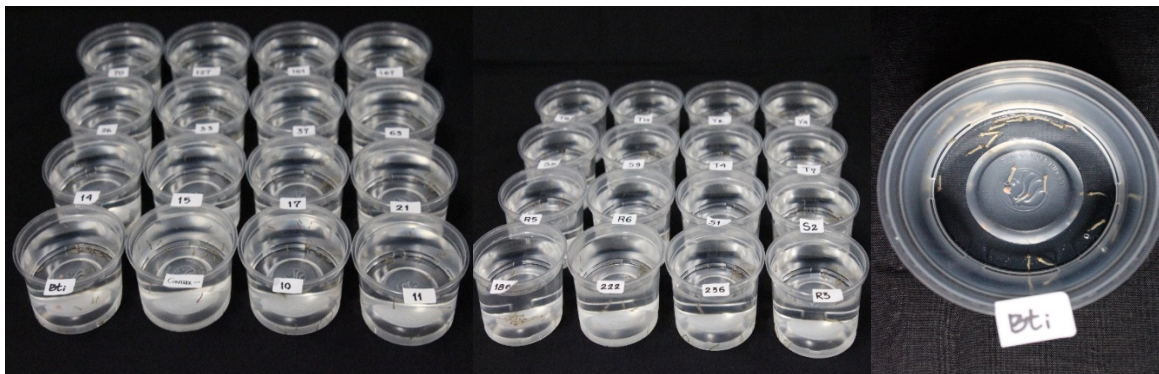
ANEXO 4

Péptidos obtenidos de la proteína de la cepa LBIC-004 y 4Q7 recombinante.

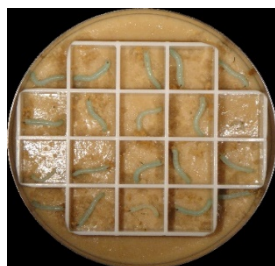
Cepa	Peso de la proteína	Péptidos (Más de 4 péptidos)	% de Covertura	Probabilidad
LBIC-004	74.9 kDa	ALAELEGLQNIMKSY LYWNNKNEKGGTIPIFGWTHHSVAPK YRREMTLAVLDIVALLPNYDIYMYPMPVHAELTR ITQVPVIKSNDBGTQSEVIQGGYTGGNL QYEYYNYWVGLVVKQLSASEPNVIITSGEITSEKDR TIDSTEELNNHTLSFIESFDLYWNNKNEKGGTIPI ITQVPVIKSNDBGTQSEVIQGGYTGGNLIR GMPQNEIDIFYNEQLNSLEKYSDHCVWEYHTGLNQLK QSLSASEPNVIITSGEITSEKD EMTLAVLDIVALLPNYDIYMYPMPVHAELTREIYTDPV LVGVNQVEFYLVNNTPEK	99.7 %	100%
4Q7 Recombinante	25 kDa	IFEDKITQVPVIK QPTFQELENLIR ALAELEGLQNIMK AADSDGLLEVQIHK EIYTDVPVGSYLPYKG QSLSASEPNVIITSGEITSEKDR VSEGAENEVLSN NIIFEDKITQVPVIKSN NYEVSLLPVYAQAANLHLLLLRDS IRGSRYMGNLTITPPSSK DIVVMQSQKTIDST	99.7%	100%

ANEXO 5

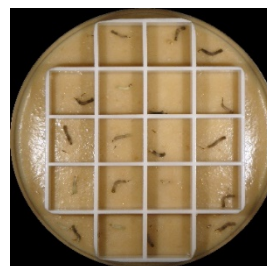
Bioensayos realizados con diferentes insectos y nemátodos



Bioensayo de las 32 cepas (LBIS10, LBIS11, LBIS14, LBIS15, LBIS17, LBIS21, LBIS26, LBIS33, LBIS37, LBIS63, LBIS70, LBIS127, LBIS161 y LBIS167, LBIS186, LBIS222, LBIS236, R3, R5, R6, S1, S2, S5, S9, T4, T7, T8, T10, Y2, Y4, CC4, IR) de *Bacillus* spp en 100mL de agua con 20 larvas de *Aedes aegypti* de 4º estadio temprano con una asada de la cepa. Duración del bioensayo 24 horas. Observaciones: Bti tuvo una mortalidad del 100% en 1 hora. Las muestras no presentaron toxicidad.



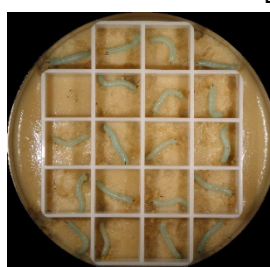
Control Negativo



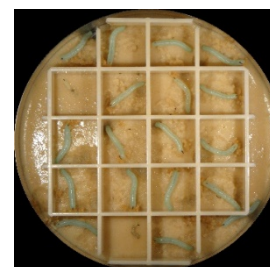
Bt-k Lisado



LBIC-004 Vegetativo



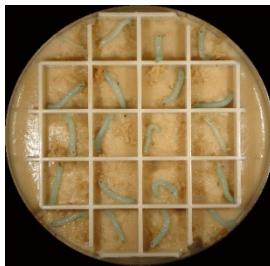
LBIC-004 S. Vegetativo



LBIC-004 Esporangio



LBIC-004 S. Esporangio

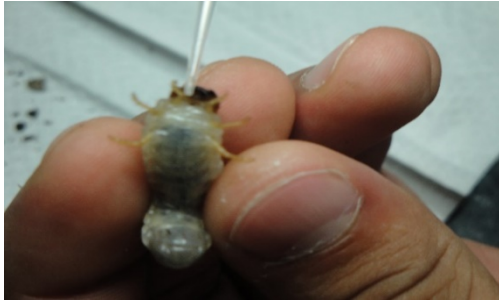
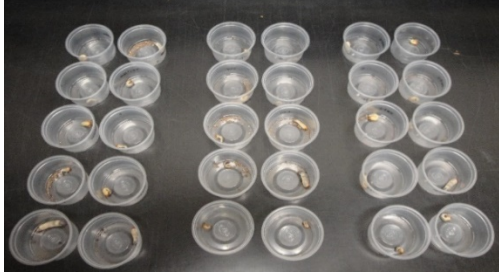


LBIC-004 Lisado



LBIC-004 S. Lisado

Bioensayo contra *M. sexta* con la cepa LBIC-004. Se utilizaron 20 larvas en placas Petri de 60cm² de dieta artificial con 10µg x cm² de cada tratamiento. No hubo toxicidad. S=Sobrenadante.



Bioensayo contra *Phyllophaga* sp., para este bioensayo se utilizaron 100µg de esporangio por Larva en 3µL de solución de Tween 80 al 0.02% y sacarosa al 20%. Se utilizaron 10 larvas por tratamiento y el bioensayo se llevó a cabo por 5 días. No hubo toxicidad.



LBIC-CC4



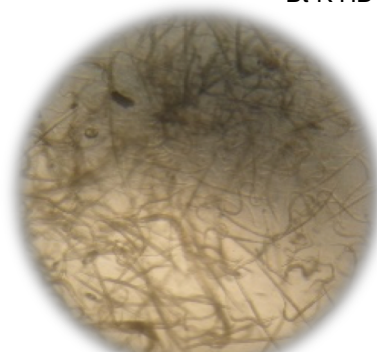
4Q7-2



Bt-K HD1



Bt-107



E. coli

Bioensayo realizado contra *C. elegans* con 100µg/mL de solución de esporangio por tratamiento. El bioensayo se realizó con más de 20 nemátodos por tratamiento (~30), con una duración del bioensayo de 7 días (Se prolongó por 9 días). No se observó mortalidad significativa.

ANEXO 6

Datos obtenidos de los 857 genomas obtenidos de la base de datos del GenBank y analizados para la búsqueda de proteínas Cry con el modelo HMM.

#	Genoma	Especie y/o cepa	Descripción	ID Proteína	E-Value	Cry
1	GCA_000161475.1	<i>Bacillus thuringiensis</i> serovar <i>tochigiensis</i> BGSC 4Y1	hypothetical protein bthur0001_58040	EEM19090.1	7.70E-178	Cry7Aa
			hypothetical protein bthur0001_54740	EEM19403.1	2.20E-171	Cry7Ea
			hypothetical protein bthur0001_55730	EEM19308.1	2.10E-134	Cry61Aa
			Pesticidal crystal protein cry4Aa	EEM19417.1	1.20E-84	Cry4Aa
2	GCA_000161515.1	<i>Bacillus thuringiensis</i> serovar <i>thuringiensis</i> str. T01001	Pesticidal crystal protein cry1Ae	EEM32076.1	0	Cry1Ae
			Pesticidal crystal protein cry1Ae	EEM32076.1	0	Cry1Ae
			Pesticidal crystal protein cry1Bc	EEM31561.1	0	Cry1Bc
3	GCA_000161535.1	<i>Bacillus thuringiensis</i> serovar <i>sotto</i> str. T04001	Pesticidal crystal protein cry1Ae	EEM38018.1	0	Cry1Ae
4	GCA_000161555.1	<i>Bacillus thuringiensis</i> serovar <i>pakistani</i> str. T13001	Cry21Ba1	EEM44457.1	7.90E-57	Cry21Ba1
			Pesticidal crystal protein cry5Ba	EEM44869.1	2.70E-52	Cry5Ba
			Pesticidal crystal protein cry5Ba	EEM44315.1	3.70E-41	Cry5Ba
5	GCA_000161575.1	<i>Bacillus thuringiensis</i> serovar <i>kurstaki</i> str. T03a001	Pesticidal crystal protein cry1Ac	EEM50110.1	0	Cry1Ac
6	GCA_000161595.1	<i>Bacillus thuringiensis</i> serovar <i>monterrey</i> BGSC 4AJ1	Pesticidal crystal protein cry7Aa	EEM55953.1	1.50E-177	Cry7Aa
7	GCA_000161615.1	<i>Bacillus thuringiensis</i> serovar <i>berliner</i> ATCC 10792	Pesticidal crystal protein cry1Ae	EEM62204.1	0	Cry1Ae
			Pesticidal crystal protein cry1Ae	EEM62204.1	0	Cry1Ae
			Pesticidal crystal protein cry1Bc	EEM62552.1	0	Cry1Bc
8	GCA_000161635.1	<i>Bacillus thuringiensis</i> serovar <i>andalousiensis</i> BGSC 4AW1	hypothetical protein bthur0009_53090	EEM68666.1	5.50E-105	Cry32Eb
			hypothetical protein bthur0009_56310	EEM68354.1	3.40E-102	Cry32Eb
9	GCA_000161695.1	<i>Bacillus thuringiensis</i> serovar <i>pulsiensis</i> BGSC 4CC1	Pesticidal crystal protein cry8Ba	EEM86551.1	5.40E-135	Cry8Ba

		hypothetical protein bthur0012_54190	EEM86553.1	2.10E-81	Cry8	
10	GCA_000161715.1	<i>Bacillus thuringiensis</i> IBL 200	Pesticidal crystal protein cry1Ae	EEM92934.1	0	Cry1Ae
		Pesticidal crystal protein cry1Ae	EEM92953.1	0	Cry1Ae	
		Pesticidal crystal protein cry1Ae	EEM92924.1	0	Cry1Ae	
		Pesticidal crystal protein cry1Ae	EEM92947.1	0	Cry1Ae	
		Pesticidal crystal protein cry1Bc	EEM92927.1	0	Cry1Bc	
11	GCA_000190515.1	<i>Bacillus thuringiensis</i> serovar finitimus YBT-020	pesticidal crystal protein cry4AA (plasmid)	ADY24870.1	4.00E-151	Cry4Aa
		pesticidal crystal protein cry4AA (plasmid)	ADY25012.1	4.00E-151	Cry4Aa	
		pesticidal crystal protein cry4BA (plasmid)	ADY24756.1	2.80E-141	Cry4Ba	
		pesticidal crystal protein cry4BA (plasmid)	ADY24720.1	8.30E-104	Cry4Ba	
12	GCA_000193355.1	<i>Bacillus thuringiensis</i> serovar chinensis CT-43	Pesticidal crystal protein cry1Aa (plasmid)	AEA19608.1	0	Cry1Aa
		Pesticidal crystal protein cry1Ba (plasmid)	AEA19203.1	0	Cry1Ba	
13	GCA_000290695.1	<i>Bacillus cereus</i> VDM022	hypothetical protein IKM_06016	EJR93120.1	8.50E-126	Cry32Da
		hypothetical protein IKM_05491	EJR94915.1	3.20E-112	Cry32Eb	
		hypothetical protein IKM_05477, partial	EJR95133.1	2.00E-109	Cry32	
		hypothetical protein IKM_05464, partial	EJR95827.1	1.80E-106	Cry32Da	
14	GCA_000290795.1	<i>Bacillus cereus</i> VD154	hypothetical protein IK5_04822	EJR69327.1	8.30E-57	Cry21Ba1
		pesticidal crystal protein cry5Aa	EJR71967.1	2.90E-52	Cry5Aa	
		pesticidal crystal protein cry5Aa	EJR70782.1	3.90E-41	Cry5Aa	
15	GCA_000292455.1	<i>Bacillus thuringiensis</i> HD-771	Pesticidal crystal protein cry1Aa (plasmid)	AFQ19845.1	0	Cry1Aa
16	GCA_000292705.1	<i>Bacillus thuringiensis</i> HD-789	pesticidal crystal protein cry4BA (plasmid)	AFQ30496.1	3.00E-135	Cry4Ba
		pesticidal crystal protein cry4AA (plasmid)	AFQ30558.1	3.00E-133	Cry4Aa	
		putative pesticidal crystal protein (plasmid)	AFQ30561.1	1.40E-124	Cry4Ba	
17	GCA_000293685.1	<i>Bacillus cereus</i> HD73	pesticidal crystal protein cry1Ac	EJV73248.1	0	Cry1Ac
18	GCA_000300475.1	<i>Bacillus thuringiensis</i> MC28	parasporal crystal protein (plasmid)	AFU17214.1	6.20E-130	Cry4Ba
		pesticidal crystal protein cry4BA (plasmid)	AFU16949.1	2.10E-126	Cry4Ba	
		pesticidal crystal protein (plasmid)	AFU17015.1	1.80E-122	Cry4Cb	
		Cry40-like protein (plasmid)	AFU17353.1	2.40E-121	Cry40	
		Cry39ORF2 protein (plasmid)	AFU17332.1	8.70E-120	Cry39	

			C-terminal half of 130-kDa type delta-endotoxin (AFU17343.1	1.90E-114	Cry39
19	GCA_000338755.1	<i>Bacillus thuringiensis</i> serovar kurstaki str. HD73	Pesticidal crystal protein cry1Ac (plasmid)	AGE81486.1	0	Cry1Ac
20	GCA_000341665.1	<i>Bacillus thuringiensis</i> serovar thuringiensis str. IS5056	Insecticidal crystal protein Cry1Aa3 (plasmid)	AGG05325.1	0	Cry1Aa3
			Insecticidal crystal protein Cry1Ab21 (plasmid)	AGG04467.1	0	Cry1Ab21
			Insecticidal crystal protein Cry1Ab21 (plasmid)	AGG04467.1	0	Cry1Ab21
			Insecticidal crystal protein Cry1Ba1 (plasmid)	AGG04729.1	0	Cry1Ba1
21	GCA_000387405.2	<i>Bacillus thuringiensis</i> T01-328	Insecticidal delta-endotoxin CryIA(a)	ERH97550.1	0	Cry1A
			Insecticidal delta-endotoxin CryIA(a)	ERH96459.1	0	Cry1A
22	GCA_000399005.1	<i>Bacillus cereus</i> BAG1X1-1	hypothetical protein ICC_06369	EOO22999.1	1.20E-130	Cry4Ba
			pesticidal crystal protein cry4Ba	EOO22988.1	1.80E-130	Cry4Ba
23	GCA_000399065.1	<i>Bacillus cereus</i> BAG1X2-1	pesticidal crystal protein cry4Ba	EOO42768.1	5.90E-131	Cry4Ba
			hypothetical protein ICI_06286	EOO42779.1	1.30E-130	Cry4Ba
24	GCA_000399085.1	<i>Bacillus cereus</i> BAG1X2-2	hypothetical protein ICK_06583	EOO43890.1	8.90E-131	Cry4Ba
			pesticidal crystal protein cry4Ba	EOO43879.1	1.90E-130	Cry4Ba
25	GCA_000399105.1	<i>Bacillus cereus</i> BAG1X2-3	hypothetical protein ICM_06222	EOO55921.1	1.20E-130	Cry4Ba
			hypothetical protein ICM_06211	EOO55910.1	1.80E-130	Cry4Ba
26	GCA_000399265.1	<i>Bacillus cereus</i> BAG2O-1	hypothetical protein ICO_06633	EOO99861.1	1.20E-130	Cry4Ba
			pesticidal crystal protein cry4Ba	EOO99989.1	1.90E-130	Cry4Ba
			hypothetical protein ICO_06632, partial	EOP00007.1	8.60E-85	Cry4Ba
27	GCA_000399345.1	<i>Bacillus cereus</i> BMG1.7	pesticidal crystal protein cry1Aa	EOP87781.1	0	Cry1Aa
			pesticidal crystal protein cry1Ac	EOP80401.1	0	Cry1Ac
28	GCA_000399365.1	<i>Bacillus cereus</i> HuB13-1	pesticidal crystal protein cry1Ac, partial	EOP28672.1	0	Cry1Ac
29	GCA_000399505.1	<i>Bacillus cereus</i> ISP2954	pesticidal crystal protein cry1Ab, partial	EOP50953.1	0	Cry1Ab
			pesticidal crystal protein cry1Ab, partial	EOP50953.1	0	Cry1Ab
			pesticidal crystal protein cry1Aa, partial	EOP61985.1	1.80E-129	Cry1Aa
			pesticidal crystal protein cry1Ac, partial	EOP61889.1	2.80E-114	Cry1Ac
30	GCA_000468975.2	<i>Bacillus thuringiensis</i> serovar aizawai str. Leapi01	hypothetical protein C621_0200655	ETE96065.1	1.80E-129	Cry9Ea
31	GCA_000468995.3	<i>Bacillus thuringiensis</i> serovar aizawai str. Hu4-2	hypothetical protein C623_0213410, partial	ETE97631.1	0	Cry9Ea
			Pesticidal crystal protein cry1Aa, partial	ETE96169.1	3.00E-306	Cry1Aa

			hypothetical protein C623_0225770	ETE93019.1	2.00E-129	Cry9Ea
32	GCA_000497525.2	<i>Bacillus thuringiensis</i> YBT-1518	pesticidal crystal protein cry5Ba	AHA70031.1	1.40E-48	Cry5Ba
33	GCA_000530375.1	<i>Bacillus thuringiensis</i> DB27	unnamed protein product, partial	CDN39788.1	5.60E-59	Cry21Ga1
			unnamed protein product	CDN39519.1	1.20E-58	Cry21Fa1
			unnamed protein product	CDN33563.1	1.80E-52	Cry21Ba1
			unnamed protein product, partial	CDN39777.1	2.00E-47	Cry21Ha1
			unnamed protein product, partial	CDN33656.1	4.30E-40	Cry21Ha1
34	GCA_000600315.1	<i>Bacillus thuringiensis</i>	pesticidal protein	EXY04520.1	0	Cry1Ga
			pesticidal protein, partial	EXY04555.1	0	Cry1Ab35
			pesticidal protein	EXY04526.1	0	Cry1Cb
			pesticidal protein, partial	EXY04468.1	0	Cry1
			pesticidal protein	EXY05451.1	0	Cry1Db
			pesticidal protein	EXY05452.1	0	Cry1F
			pesticidal protein	EXY04803.1	0	Cry1Bd
35	GCA_000688795.1	<i>Bacillus thuringiensis</i> serovar kurstaki str. YBT-1520	Pesticidal crystal protein cry1Aa (plasmid)	AHZ54991.1	0	Cry1Aa
			Pesticidal crystal protein cry1Aa (plasmid)	AHZ55258.1	0	Cry1Aa
36	GCA_000710255.1	<i>Bacillus thuringiensis</i> serovar kurstaki str. HD-1	hypothetical protein BG09_2546	KEH48722.1	0	Cry1A
			hypothetical protein BG09_6053	KEH45261.1	0	Cry1A
			hypothetical protein BG09_6053	KEH45261.1	0	Cry1A
37	GCA_000717535.1	<i>Bacillus thuringiensis</i> serovar kurstaki str. HD-1	Pesticidal crystal protein cry1Aa (plasmid)	AIE37439.1	0	Cry1Aa
			Pesticidal crystal protein cry1Aa (plasmid)	AIE37211.1	0	Cry1Aa
			Pesticidal crystal protein cry1Aa (plasmid)	AIE37011.1	0	Cry1Aa
			Pesticidal crystal protein cry1Aa (plasmid)	AIE37011.1	0	Cry1Aa
38	GCA_000747545.1	<i>Bacillus thuringiensis</i> serovar kurstaki str. YBT-1520	dipterans toxin crystal protein Cry1Aa3 (plasmid)	AIM34741.1	0	Cry1Aa3
			dipterans toxic crystal protein-like protein Cry1	AIM34400.1	0	Cry1Ac
			dipterans toxic crystal protein-like protein Cry1	AIM34414.1	0	Cry1Ac
39	GCA_000774075.2	<i>Bacillus thuringiensis</i>	pesticidal protein (plasmid)	ALL11724.1	1.50E-180	Cry7B
			pesticidal protein	ALL21863.1	1.50E-180	Cry7B
			pesticidal protein (plasmid)	ALL11841.1	7.40E-178	Cry7Aa

		pesticidal protein	ALL21979.1	7.40E-178	Cry7Aa	
		hypothetical protein BTXL6_28185 (plasmid)	ALL11744.1	2.50E-177	Cry7	
		hypothetical protein BTXL6_10620	ALL21883.1	2.50E-177	Cry7	
		pesticidal protein (plasmid)	ALL11871.1	2.30E-175	Cry7Ca	
		pesticidal protein	ALL25005.1	2.30E-175	Cry7Ca	
		hypothetical protein BTXL6_28705 (plasmid)	ALL11828.1	2.10E-171	Cry7Ea	
		hypothetical protein BTXL6_11140	ALL21966.1	2.10E-171	Cry7Ea	
		hypothetical protein BTXL6_28855 (plasmid)	ALL11846.1	2.00E-134	Cry61Aa	
		hypothetical protein BTXL6_11290	ALL21984.1	2.00E-134	Cry61Aa	
		hypothetical protein BTXL6_28770 (plasmid)	ALL11835.1	1.10E-84	Cry4Aa	
		hypothetical protein BTXL6_11205	ALL21973.1	1.10E-84	Cry4Aa	
40	GCA_000803665.1	<i>Bacillus thuringiensis</i> serovar galleriae	pesticidal protein (plasmid)	AJA23506.1	0	Cry1Aa
			pesticidal protein (plasmid)	AJA23086.1	0	Cry1Ca
			pesticidal protein (plasmid)	AJA23088.1	0	Cry1Da
			pesticidal protein (plasmid)	AJA23643.1	0	Cry1Ac
			hypothetical protein BT4G5_32525 (plasmid)	AJA23511.1	1.90E-129	Cry9Ea
41	GCA_000833655.1	<i>Bacillus thuringiensis</i> serovar morrisoni	delta endotoxin family protein	KIP25650.1	0	Cry1Bd
			pesticidal crystal protein cry1Aa	KIP24267.1	0	Cry1Aa
42	GCA_000835235.1	<i>Bacillus thuringiensis</i> serovar kurstaki	pesticidal crystal protein cry1Aa (plasmid)	AJK37978.1	0	Cry1Aa
			pesticidal crystal protein cry1Ac (plasmid)	AJK38183.1	0	Cry1Ac
			pesticidal crystal protein cry1Ab (plasmid)	AJK38600.1	0	Cry1Ab
			pesticidal crystal protein cry1Ab (plasmid)	AJK38600.1	0	Cry1Ab
43	GCA_000836715.1	<i>Bacillus</i> sp. L_1B0_5	hypothetical protein RT27_24655, partial	KIQ82002.1	4.30E-59	Cry4Aa
			pesticidal crystal protein cry4Aa, partial	KIQ78200.1	3.30E-54	Cry4Aa
			pesticidal crystal protein, partial	KIQ76838.1	8.00E-53	Cry40
44	GCA_000836745.1	<i>Bacillus</i> sp. L_1B0_8	hypothetical protein RW25_30125, partial	KIQ76870.1	4.10E-59	Cry4Aa
			pesticidal crystal protein, partial	KIQ80416.1	7.50E-53	Cry40
45	GCA_000878525.1	<i>Bacillus thuringiensis</i> Sbt003	pesticidal crystal protein cry4BA	KIU76093.1	8.40E-57	Cry4Ba
			pesticidal crystal protein cry4AA	KIU70667.1	2.90E-52	Cry4Aa

			pesticidal crystal protein cry4AA	KIU74325.1	4.00E-41	Cry4Aa
46	GCA_000969665.1	<i>Bacillus thuringiensis</i> serovar mexicanensis	Pesticidal crystal protein cry1Ac	KKB28329.1	0	Cry1Ac
47	GCA_001008665.1	<i>Bacillus cereus</i>	hypothetical protein B4158_6152	KLA17702.1	1.80E-129	Cry9Ea
48	GCA_001017635.1	<i>Bacillus thuringiensis</i>	pesticidal protein (plasmid)	AKJ62790.1	0	Cry1A
			pesticidal protein (plasmid)	AKJ62760.1	0	Cry1Ac
49	GCA_001044855.1	<i>Bacillus cereus</i>	pesticidal protein	KMQ19863.1	0	Cry1Aa
			pesticidal protein	KMQ19869.1	0	Cry1Ca
			pesticidal protein	KMQ19871.1	0	Cry1Da
			hypothetical protein TU66_09770	KMQ16968.1	1.90E-129	Cry9Ea
50	GCA_001182785.1	<i>Bacillus thuringiensis</i>	pesticidal protein (plasmid)	AKR13097.1	6.70E-126	Cry4Ba
			pesticidal crystal protein (plasmid)	AKR12910.1	8.00E-120	Cry39
			pesticidal crystal protein (plasmid)	AKR12900.1	4.00E-118	Cry39
			pesticidal crystal protein (plasmid)	AKR13120.1	1.30E-115	Cry4Cb
			pesticidal crystal protein (plasmid)	AKR13105.1	9.10E-115	Cry39
			pesticidal crystal protein (plasmid)	AKR13112.1	3.20E-114	Cry39
			pesticidal crystal protein (plasmid)	AKR13281.1	1.50E-111	Cry39
			pesticidal crystal protein (plasmid)	AKR13092.1	3.60E-110	Cry39
51	GCA_001183785.1	<i>Bacillus thuringiensis</i> serovar indiana	Pesticidal crystal protein cry7Ab (plasmid)	AKR38949.1	5.80E-179	Cry7Ab
			Pesticidal crystal protein cry7Aa (plasmid)	AKR38911.1	1.60E-178	Cry7Aa
			Pesticidal crystal protein cry7Aa (plasmid)	AKR38906.1	8.00E-176	Cry7Aa
52	GCA_001286785.1	<i>Bacillus subtilis</i>	Pesticidal crystal protein cry14Aa	CUB55363.1	6.70E-37	Cry14Aa
53	GCA_001402735.1	<i>Bacillus thuringiensis</i>	pesticidal crystal protein (plasmid)	KQB17993.1	1.10E-124	Cry4Ba
			pesticidal crystal protein cry4Aa, partial (plasmid)	KQB18012.1	1.50E-50	Cry4Aa
54	GCA_001420855.1	<i>Bacillus thuringiensis</i>	pesticidal protein (plasmid)	ALL62164.1	0	Cry1Ac
55	GCA_001548175.1	<i>Bacillus thuringiensis</i> serovar tolworthi	pesticidal crystal protein cry1Aa (plasmid)	BAR87554.1	0	Cry1Aa
56	GCA_001583935.1	<i>Bacillus cereus</i>	hypothetical protein AT268_27915, partial	KXY40528.1	1.60E-143	Cry8Aa
			hypothetical protein AT268_15015	KXY38626.1	9.80E-84	Cry8
57	GCA_001583975.1	<i>Bacillus cereus</i>	hypothetical protein AT270_10220	KXY85197.1	6.50E-58	Cry21fa
			hypothetical protein AT270_09435	KXY85247.1	5.00E-54	Cry21Ba

58	GCA_001595725.1	<i>Bacillus thuringiensis</i>	hypothetical protein AT270_30505, partial	KXY84416.1	4.20E-50	Cry21Fa1
			hypothetical protein AXW78_30240 (plasmid)	AMR06122.1	1.90E-140	Cry8Ea
			hypothetical protein AXW78_30115 (plasmid)	AMR06312.1	1.00E-139	Cry8X
59	GCA_001598095.1	<i>Bacillus thuringiensis</i>	pesticidal protein (plasmid)	AMR88565.1	0	Cry1Ab35
			pesticidal protein (plasmid)	AMR88578.1	0	Cry1Da
			pesticidal protein (plasmid)	AMR88562.1	0	Cry1H
			pesticidal protein (plasmid)	AMR88428.1	0	Cry1Fb
			pesticidal protein (plasmid)	AMR88574.1	0	Cry1Ja
			pesticidal protein (plasmid)	AMR88563.1	0	Cry1Bb
60	GCA_001640965.1	<i>Bacillus thuringiensis</i> serovar alesti	pesticidal protein (plasmid)	AND11141.1	0	Cry1Ae

ANEXO 7

Información sobre algunas cepas de *Bacillus* spp. utilizadas en el proyecto, su origen y características fenotípicas.

Número de muestra	Fecha	Región	Sitio de aislamiento	Profundidad	Fenotipo
20a	oct-07	Churince	LI old 1	Sedimento top	Naranja
33	oct-07	Churince	LI old 1	Sedimento top	Naranja
34	oct-07	Churince	LI old 1	Sedimento top	Naranja
35	oct-07	Churince	LI old 1	Sedimento top	Salmón
36	oct-07	Churince	LI old 1	Sedimento top	Naranja cremosa, brillantes, colonias chiquitas
37	oct-07	Churince	LI old 1	Sedimento top	Salmón, cremosa opacas, colonias chiquitas
43	oct-07	Churince	LI old 1	Sedimento top	Rosa pálido, algodonosa, opaca con costras
48	oct-07	Churince	LI old 2	B	Beige translúcida
73b	oct-07	Churince	LI old 2	B	Naranja
81a	oct-07	Churince	LI old 3	Sedimento top	Naranja claro semitranslúcida
81b	oct-07	Churince	LI old 3	Sedimento top	Naranja claro semitranslúcida
90	oct-07	Churince	LI old 3	Sedimento top	Beige
91b	oct-07	Churince	LI old 3	Sedimento top	Mostaza, brillante semitranslúcida como con un ombligo
95b	oct-07	Churince	LI old 3	Sedimento top	Beige algodonosa
98b	oct-07	Churince	LI old 3	Sedimento top	Amarilla, brillante, lisa, colonias redondas
101a	oct-07	Churince	LI old 3	Sedimento top	Blanca lisa brillante
102	oct-07	Churince	LI old 3	Sedimento top	Blanca lisa brillante (colonias como nubes)
109a	oct-07	Churince	LI old 4	Sedimento top	Blanca lisa brillante
109b	oct-07	Churince	LI old 4	Sedimento top	Blanca lisa brillante
110	oct-07	Churince	LI old 4	Sedimento top	Crema rugosa (forma costras)
111b	oct-07	Churince	LI old 4	Sedimento top	Beige, algodonosa, asimétrica y opaca
112a	oct-07	Churince	LI old 4	Sedimento top	Blanca opaca redonda
112b	oct-07	Churince	LI old 4	Sedimento top	Naranja lisa brillante
113a	oct-07	Churince	LI old 4	Sedimento top	Naranja pálido lisa brillante
113b	oct-07	Churince	LI old 4	Sedimento top	Naranja pálido lisa brillante
120a	oct-07	Churince	LI old 4	Sedimento deep	Rosa pálido orillas algodonosas opaca
120b	oct-07	Churince	LI old 4	Sedimento deep	Blanca opaca orillas algodonosas
121a	oct-07	Churince	LI old 4	Sedimento deep	Blanca

122a	oct-07	Churince	LI old 4	Sedimento deep	Blanca lisa asimétrica
122b	oct-07	Churince	LI old 4	Sedimento deep	Crema lisa asimétrica
123a	oct-07	Churince	LI old 4	Sedimento deep	Naranja Lisa Brillante redonda
124a	oct-07	Churince	LI old 4	Sedimento deep	Blanca algodonosa orillas asimétricas/Rosa algodonosa
126	oct-07	Churince	LI old 4	Sedimento deep	Blanca rugosa algodonosa
127a	oct-07	Churince	LI old 4	Sedimento deep	Naranja lisa redonda brillante grande
127b	oct-07	Churince	LI old 4	Sedimento deep	Naranja claro lisa redonda Brillante
128a	oct-07	Churince	LI old 4	Sedimento deep	Rosa pálido lisa brillante redonda
128b	oct-07	Churince	LI old 4	Sedimento deep	Crema fuerte lisa brillante
130a	oct-07	Churince	LI old 4	Sedimento deep	Rosa algodonosa/Blanca translúcida alrededor rosa
130b	oct-07	Churince	LI old 4	Sedimento deep	Rosa algodonosa/Blanca translúcida alrededor rosa
131a	oct-07	Churince	LI old 4	Sedimento deep	Rosa algodonosa/Blanca translúcida alrededor rosa
132	oct-07	Churince	LI old 4	Sedimento deep	Rosa rugosa brillante orillas asimétricas
133	oct-07	Churince	LI old 4	Sedimento deep	Rosa rugosa brillante orillas asimétricas
134b	oct-07	Churince	LI old 4	Sedimento deep	Forma de ombligo tiene una colonia crema translúcida alrededor
134c	oct-07	Churince	LI old 4	Sedimento deep	Rosa pálido lisa brillante redonda
135a	oct-07	Churince	LI old 4	Sedimento top	Naranja brillante lisa redonda
135b	oct-07	Churince	LI old 4	Sedimento top	Amarilla lisa brillante redonda
136a	oct-07	Churince	LI old 4	Sedimento top	Amarilla redonda lisa
138	oct-07	Churince	LI old 4	Sedimento top	Crema lisa brillante
139b	oct-07	Churince	LI old 4	Sedimento top	Blanca algodonosa orillas asimétricas
140a	oct-07	Churince	LI old 4	Sedimento top	Blanca brillante asimétrica forma costras
141b	oct-07	Churince	LI old 4	Sedimento top	Naranja lisa brillante redonda
142	oct-07	Churince	LI old 4	Sedimento top	Rosa orillas asimétricas
143	oct-07	Churince	LI old 4	Sedimento top	Amarilla lisa brillante redonda
144a	oct-07	Churince	LI old 4	Sedimento top	Blanca lisa brillante
144b	oct-07	Churince	LI old 4	Sedimento top	Crema redonda brillante
145	oct-07	Churince	LI old 4	Sedimento top	Crema lisa brillante
146a	oct-07	Churince	LI old 4	Sedimento top	Crema redonda pequeña
146b	oct-07	Churince	LI old 4	Sedimento top	Crema translúcida lisa redonda más grande
147a	oct-07	Churince	LI old 4	Sedimento top	Amarilla lisa brillante
150a	oct-07	Churince	LI old 4	Sedimento top	Crema claro lisa brillante abultada
150ba	oct-07	Churince	LI old 4	Sedimento top	Crema brillante lisa

150bb	oct-07	Churince	LI old 4	Sedimento top	Crema brillante lisa
155b (LBIC-T7)	oct-07	Churince	LI old 5	Sedimento top	Blanca lisa brillante colonias grandes asimétrica
156	oct-07	Churince	LI old 5	Sedimento top	NE
158b	oct-07	Churince	LI old 5	Sedimento top	Amarilla brillante redonda
159a	oct-07	Churince	LI old 5	Sedimento top	Amarilla lisa brillante
160b	oct-07	Churince	LI old 5	Sedimento top	Blanca lisa brillante orillas asimétricas (algodonosa)
162	oct-07	Churince	LI old 5	Sedimento top	Rosa pálido lisa brillante
163b	oct-07	Churince	LI old 5	Sedimento top	Crema lisa brillante
183	oct-07	Churince	LI old 5	Sedimento deep	Blanca rosa forma dedos
311a	oct-07	Churince	LI old 11	Sedimento top	Blanca rugosa opaca, colonias redondas
311b	oct-07	Churince	LI old 11	Sedimento top	Blanca opaca algodonosa
313b	oct-07	Churince	LI old 11	Sedimento top	Amarillo lisa brillante
316	oct-07	Churince	LI old 11	Sedimento top	Blanco algodonoso opaca
319b	oct-07	Churince	LI old 11	Sedimento top	Blanco algodonoso orillas asimétrica
321	oct-07	Churince	LI old 11	Sedimento top	Blanco algodonoso colonias amorfas rugosas brillantes
325b	oct-07	Churince	LI old 11	C	Blanca lisa brillante abultada como gotitas de agua
327a	oct-07	Churince	LI old 11	Sedimento deep	Rosa/Blanco algodonoso
339	oct-07	Churince	LI old 11	Sedimento deep	Rosa/Blanco algodonoso
357	oct-07	Churince	LI old 12	Sedimento top	Naranja fuerte lisa brillante colonias redondas medianas
359a	oct-07	Churince	LI old 12	Sedimento top	Blanca
370a	oct-07	Churince	LI old 12	Sedimento top	Blanca algodonosa opaca colonias asimétricas grandes
370b	oct-07	Churince	LI old 12	Sedimento top	Algodonosa orillas peludas
372	oct-07	Churince	LI old 12	Sedimento deep	Blanca algodonoso
416a (LBIC-Y4)	oct-07	Churince	LI old 13	Sedimento top	Crema algodonosa colonias grandes algo rugosas
416c	oct-07	Churince	LI old 13	Sedimento top	Naranja translúcida
417	oct-07	Churince	LI old 13	Sedimento top	Crema algodonosa colonias grandes algo rugosas en las orillas
418b	oct-07	Churince	LI old 13	Sedimento top	Naranja translúcida
419a	oct-07	Churince	LI old 13	Sedimento top	Naranja opaco
419b	oct-07	Churince	LI old 13	Sedimento top	Naranja brillante
421b	oct-07	Churince	LI old 13	Sedimento top	Crema opaca colonias pequeñas juntas son rugosas en las orillas
422a	oct-07	Churince	LI old 13	Sedimento top	Blanca
423b	oct-07	Churince	LI old 13	Sedimento top	Naranja translúcido colonias grandes sin contorno definido

424	oct-07	Churince	LI old 13	Sedimento top	algo rugosas en orillas NE
425b	oct-07	Churince	LI old 13	Sedimento top	Blanca algodonosa colonias grandes redondas y rugosas
425c	oct-07	Churince	LI old 13	Sedimento top	Salmón brillante colonias medianas redondas lisas en las orillas solamente
426a (LBIC-T4)	oct-07	Churince	LI old 13	Sedimento top	Crema opaca algodonosa colonias grandes que no se separan
427	oct-07	Churince	LI old 13	Sedimento top	Rosa/Blanco peludita colonias medianas no muy separadas y rugosas
428a (LBIC-R6)	oct-07	Churince	LI old 13	Sedimento top	Crema algodonosa de colonias medianas no bien separadas rugosas en las orillas
429b	oct-07	Churince	LI old 13	Sedimento top	Crema brillante colonias pequeñas juntas y rugosas al centro
430a	oct-07	Churince	LI old 13	Sedimento deep	Naranja translúcido
431a	oct-07	Churince	LI old 13	Sedimento deep	Crema brillantes colonias pequeñas rugosas y sin forma definida
435a	oct-07	Churince	LI old 13	Sedimento deep	Beige redonada y brillante con colonias muy pequeñas
436a	oct-07	Churince	LI old 13	Sedimento deep	Naranja opaco con un punto de intensidad en el centro no son redondas y colonias grandes
436b	oct-07	Churince	LI old 13	Sedimento deep	Crema algodonosa de colonias muy grandes rugosas en las orillas
437b	oct-07	Churince	LI old 13	Sedimento deep	Crema translúcido colonias pequeñas sin forma definida lisas
438b	oct-07	Churince	LI old 13	Sedimento deep	Crema colonias muy grandes sin forma poco rugosas en el centro
438c	oct-07	Churince	LI old 13	Sedimento deep	Naranja translúcido colonias medianas redondas y lisas
440b	oct-07	Churince	LI old 13	Sedimento deep	Blancas brillante colonias pequeñas sin forma algo rugosas
441c	oct-07	Churince	LI old 13	Sedimento deep	Blancas colonias medianas no redondas son lisas
442a	oct-07	Churince	LI old 13	Sedimento deep	Crema colonias medianas sin forma tienen un punto un poco más fuerte en el centro
442b	oct-07	Churince	LI old 13	Sedimento deep	Crema brillante colonias pequeñas redondas y lisas
446	oct-07	Churince	LI old 14	Sedimento top	Colonias de color crema, algodonosas, pequeñas y cremosas
447	oct-07	Churince	LI old 14	Sedimento top	Beige, colonias cremosa, opacas y pequeñas
448b	oct-07	Churince	LI old 14	Sedimento top	Crema blanquecino, colonias pequeñas opacas
449a2	oct-07	Churince	LI old 14	Sedimento top	Naranja transparentes, colonias pequeñas y cremosas
451a	oct-07	Churince	LI old 14	Sedimento top	Anaranjadas brillantes cremosas (descripción agregada por Rocío) Son translúcidas en comparación con 451b
458a	oct-07	Churince	LI old 14	Sedimento deep	Crema

458b	oct-07	Churince	LI old 14	Sedimento deep	Beige cremosa, opaca, algodonosa
460b (LBIC-S9, LBIC-004)	oct-07	Churince	LI old 14	Sedimento deep	Colonias crema, cremosas, medianas, brillosas

NE: No existe, LI: Laguna intermedia, **Negritas: Cepas positivas para genes cry.** Información adicional consultar: doi: 10.3389/fmicb.2017.00029