UNIVERSIDADE FEDERAL DO TOCANTINS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIODIVERSIDADE
E BIOTECNOLOGIA - REDE BIONORTE

CARACTERIZAÇÃO DE ISOLADOS DE *Bacillus thuringiensis* COM POTENCIAL INSETICIDA: ABORDAGENS GENÔMICA E PROTEÔMICA

GISELLY BATISTA ALVES

PALMAS – TO
FEVEREIRO/2023

GISELLY BATISTA ALVES

CARACTERIZAÇÃO DE ISOLADOS DE *Bacillus thuringiensis* COM POTENCIAL INSETICIDA: ABORDAGENS GENÔMICA E PROTEÔMICA

Tese apresentada ao Curso de Doutorado do Programa de Pós-Graduação em Biodiversidade e Biotecnologia – Rede BIONORTE, na Universidade Federal do Tocantins, como requisito parcial para a obtenção do Título de Doutora em Biodiversidade e Biotecnologia.

Orientador Prof. Dr. Raimundo Wagner de Souza Aguiar

Palmas – TO

FEVEREIRO/2023

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da Universidade Federal do Tocantins

A474c Alves , Giselly Batista.

Caracterização de isolados de Bacillus thuringiensis com potencial inseticida: abordagens genômica e proteômica. / Giselly Batista Alves . – Palmas, TO, 2023.

160 f.

Tese (Doutorado) - Universidade Federal do Tocantins – Câmpus Universitário de Palmas - Curso de Pós-Graduação (Doutorado) em Biodiversidade e Biotecnologia, 2023.

Orientador: Raimundo Wagner de Souza Aguiar

1. Bacillus thuringiensis. 2. Genes cry. 3. Diptera. 4. Lepidoptera.. I. Título

CDD 660.6

Elaborado pelo sistema de geração automática de ficha catalográfica da UFT com os dados fornecidos pelo(a) autor(a).

GISELLY BATISTA ALVES

CARACTERIZAÇÃO DE ISOLADOS DE *Bacillus thuringiensis* COM POTENCIAL INSETICIDA: ABORDAGENS GENÔMICA E PROTEÔMICA

Tese apresentada ao Curso de Doutorado do Programa de Pós-Graduação em Biodiversidade e Biotecnologia – Rede BIONORTE, na Universidade Federal do Tocantins, como requisito parcial para a obtenção do Título de Doutora em Biodiversidade e Biotecnologia.

Orientador Prof. Dr. Raimundo Wagner de Souza Aguiar

Aprovada em 28/02/2023

Banca examinadora

Documento assinado digitalmente
gov.br RAIMUNDO WAGNER DE SOUZA AGUIAR
Data: 01/06/2023 11:52:14-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Raimundo Wagner de Souza Aguiar (Orientador)
Universidade Federal do Tocantins

Documento assinado digitalmente
gov.br GIL RODRIGUES DOS SANTOS
Data: 31/05/2023 21:13:42-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Gil Rodrigues dos Santos
Universidade Federal do Tocantins

Documento assinado digitalmente
gov.br KHALID HADDI
Data: 01/06/2023 08:10:17-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Khalid Haddi
Universidade Federal de Lavras

Documento assinado digitalmente
gov.br EUGENIO EDUARDO DE OLIVEIRA
Data: 02/06/2023 17:37:22-0300
Verifique em https://validar.iti.gov.br

Prof. Dr. Eugênio Eduardo de Oliveira
Universidade Federal de Viçosa

Firmado electrónicamente por:
LUIS OSWALDO VITERI JUMBO

Prof. Dr. Luis Oswaldo Viteri
Universidad Nacional de Loja

Dedico este trabalho a Maria Lucivânia, Rafael Sobczak, Patrícia Verdugo, Maira Obando, Tayane Alves e Antônio Pascoal

AGRADECIMENTOS

Ao Programa de Pós-Graduação em Biotecnologia e Biodiversidade da Bionorte por todo o apoio recebido.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa de estudos.

Ao Laboratório de Biologia Molecular da Universidade Federal do Tocantins e ao Laboratório de Virologia da Universade de Brasília por toda a infraestrutura para a realização deste trabalho.

Agradeço ao meu orientador, Prof. Dr. Raimundo Wagner de Souza Aguiar, pela orientação e oportunidade de executar este projeto.

Agradeço aos professores doutores Khalid Haddi, Gil Rodrigues dos Santos, Luis Oswaldo Viteri e Eugênio Eduardo por terem aceitado participar da banca examinadora e colaborado com o término deste trabalho.

Agradeço ao Eugênio Eduardo, Khalid Haddi, Bergmann Moares, Gil Rodrigues dos Santos, Luis Oswaldo Viteri, Marcio Akio Ootani, Manoel Mota dos Santos, Timóteo S. O. Lemes; Eliseu J. G. Pereira, Juan L. Jurat-Fuentes, Guy Smagghe, Roberto F. T. Correia, Fernando L. Melo, Marcelo Leite dias, Lara T. M. Costa, Fabrício S. Campos e Sergio D. Ascêncio que contribuíram no desenvolvimento dos artigos desta tese.

Agradeço a toda equipe do laboratório de Biologia Molecular.

Aos docentes do Programa de Pós-Graduação em Biotecnologia e Biodiversidade da Bionorte por todo conhecimento transmitido.

Agradeço aos discentes da turma 2018/1 de doutorado da Rede Bionorte do Tocantins;

E a todos que de alguma forma contribuiram para a execução deste trabalho.

RESUMO

A busca por alternativas sustentáveis que possam substituir os inseticidas químicos constitui uma importante estratégia para garantir o controle de insetos-praga de forma segura. Neste sentido, *Bacillus thuringiensis* (*Bt*) se destaca pela produção de proteínas inseticidas capazes de atuar no controle biológico de insetos-praga de diferentes ordens. Neste trabalho, buscou-se alternativas para o controle de insetos dípteros e lepidópteros a partir de bioensaios realizados com cepas de *Bt* isolados no Tocantins. Para explorar todos os genes relacionados com a patogenicidade dos isolados, os genomas dessas bactérias foram sequenciados. Além disso, nos capítulos dois e quatro, a proteômica foi integrada à análise genômica para a análise de proteínas expressas na mistura de esporos-cristais. No capítulo 1, a análise comparativa de quatro genomas de isolados de *Bacillus thuringiensis* subsp. *israelensis*, com toxicidade para *Aedes aegypti* e *Culex quinquefasciatus*, revelou alta identidade de sequência de nucleotídeos (>98%), mesmo perfil plasmidial e de proteínas pesticidas (*cry4Ba*, *cry4Aa*, *cry11Aa*, *cry10Aa*, *cyt1Aa*, *cyt2Ba* e *cytCa*). O genoma do isolado *Bt* TOD651, apresentado no capítulo 2, com atividade tóxica para *A. aegypti* e *C. quinquefasciatus* (CL_{50} de 0.011 e 0.023 µg/mL, respectivamente), apresentou regiões CDS altamente homólogas com os genes *cry11Aa3*, *cry10Aa4*, *cry4Aa4*, *cry4Ba5*, *cyt1Aa5*, *cyt1Ca1*, *cyt2Ba13*, *mpp60Aa3* e *mpp60Ba3*. A expressão das proteínas Cry11Aa3, Cry10Aa4, Cry4Aa4, Cry4Ba5, Cyt1Aa5, Cyt1Ca1, Cyt2Ba13 e Mpp60Ba3 foi identificada na mistura de esporos-cristais, em que Cry4Ba5 foi mais abundante que Cyt1Aa5. Além disso, a expressão da enzima Mppe foi a mais abundante dentre as proteases. Já no capítulo 3, a cepa *Bt* UFT038, testada para diferentes pragas de soja, apresentou maior toxicidade para *Spodoptera cosmioides* (CL_{50}=6,8 10^6/cm^2) e sua análise genômica revelou a presença dos genes *cry1Aa8*, *cry1Ac11*, *cry1Ia44*, *cry2Aa9*, *cry2Ab35* e *vip3Af5*. Por fim, no capítulo 4, a cepa *Bt* TOL651, filogeneticamente próximo a subespécie *kenyae*, foi mais tóxico para *Anticarsia gemmatalis* (LC_{50} =1.45 ng.cm^{-2}) em comparação a *Diatraea saccharalis* (LC_{50} = 73.77 ng.cm^{-2}). Sua análise genômica permitiu detectar os genes *cry1Aa18*, *cry1Ia44*, *cry2Aa9* e *cry1Ac5*, enquanto a proteômica indicou expressão das proteínas Cry1Aa18, Cry1Ac5 e Cry2Aa9, do qual a Cry1Ac5 foi mais abundante. Além disso, o fator de virulência InhA1 foi detectado e, portanto, também deve contribuir com a toxicidade deste isolado. Por fim, os isolados de *Bt* deste estudo são alternativas para o controle biológico e a caracterização genômica e genômica-proteômica são etapas importantes que poderão contribuir para o desenvolvimento de novas estratégias de biocontrole de mosquitos vetores de doenças e pragas agrícolas.

Palavras-chave: *Bacillus thuringiensis*; Genes *cry*; Diptera; Lepidoptera.

ABSTRACT

Sustainable alternatives that can replace chemical insecticides are necessary for insect pest control. Bacillus thuringiensis (*Bt*) is a bacterium that produces insecticidal proteins toxic against insect pests of different orders. In this work, we performed bioassays using *Bt* strains isolated from Tocantins against dipterous and lepidopteran insects. To explore all the genes related to the pathogenicity of the isolates, the genomes of these bacteria were sequenced. Furthermore, in chapters two and four, proteomic was combined with genomic analysis to detect proteins expressed in the spore-crystal mixture. In chapter one, a comparative analysis of four genomes of *Bacillus thuringiensis* subsp. *israelensis* strains with toxicity to *Aedes aegypti* and *Culex quinquefasciatus* revealed high nucleotide sequence identity (>98%), the same plasmids profile, and equal pesticidal protein content (*cry4Ba, cry4Aa, cry11Aa, cry10Aa, cyt1Aa, cyt2Ba,* and *cytCa*). The genome of the *Bt* TOD651 strain, presented in Chapter 2, with toxic activity to *A. aegypti* and *C. quinquefasciatus* (CL50 of 0.011 and 0.023 µg/mL, respectively), showed CDS regions highly homologous to *cry11Aa3, cry10Aa4, cry4Aa4, cry4Ba5, cyt1Aa5, cyt1Ca1, cyt2Ba13, mpp60Aa3,* and *mpp60Ba3* genes. The expression of Cry11Aa3, Cry10Aa4, Cry4Aa4, Cry4Ba5, Cyt1Aa5, Cyt1Ca1, Cyt2Ba13, and Mpp60Ba3 proteins was identified in the spore-crystal mixture, of which Cry4Ba5 was more abundant than Cyt1Aa5. The expression of the enzyme Mppe was the most abundant among the proteases. In chapter 3, the *Bt* UFT038, tested for different soybean pests, showed higher toxicity to *Spodoptera cosmioides* (CL_{50}=6.8 10^6/cm^2), and its genomic analysis revealed the presence of cry1Aa8, cry1Ac11, cry1Ia44, cry2Aa9, cry2Ab35, and vip3Af5 genes. Finally, in chapter 4, Bt strain TOL651, phylogenetically close to subspecies kenyae, was more toxic to Anticarsia gemmatalis (LC_{50} =1.45 ng.cm^{-2}) compared to Diatraea saccharalis (LC_{50} = 73.77 ng.cm^{-2}). Its genomic analysis allowed the detection of *cry1Aa18, cry1Ia44, cry2Aa9,* and *cry1Ac5* genes, while proteomics indicated expression of Cry1Aa18, Cry1Ac5, and Cry2Aa9 proteins, of which Cry1Ac5 was most abundant. In addition, the virulence factor InhA1 was detected and thus should also contribute to the toxicity of this isolate. In conclusion, the *Bt* isolates of this study are alternatives for biological control, and genomic and genomic-proteomic characterization are important steps that could contribute to the development of new biocontrol strategies for disease vector mosquitoes and agricultural pests.

Keywords: *Bacillus thuringiensis*; *cry* genes; Diptera; Lepdoptera.

SUMÁRIO

1 INTRODUCTION

Bacillus thuringiensis (*Bt*) is a Gram-positive and spore-forming bacterium that produces insecticidal proteins, such as Cry and Cyt crystalline proteins during the sporulation and Vip proteins during the vegetative phase (POHARE; WAGH; UDAYASURIYAN, 2020; HADI et al., 2020). These protein types comprise a wide variety of toxins that presents insecticidal activity against different orders of pest insects (TORRES-QUINTERO et al., 2018; BARANEK et al., 2020).

The insects of Diptera and Lepidoptera orders comprise mosquitoes vectors of human diseases and caterpillars that affect crops of agricultural, respectively, and their biological control using *Bt* has been applied as an alternative to chemical insecticides (EVANGELISTA et al., 2020; MEJIAS et al., 2020). However, insect populations have development resistant due to repeated applications of *Bt* sprays and the adoption of transgenic crops that produce Cry proteins (*Bt*-crops) without non-*Bt* refuge areas (CASTRO et al., 2019). The isolation of *Bt* strains has been part of the strategy to formulate new bioinsecticides and screen new cry genes (CASTRO et al., 2019; KAYAM et al., 2020).

The genome sequencing of *Bt* has allowed the easy detection of pesticidal genes and other genes related to pathogenicities such as virulence factors and other secondary metabolites (FAYAD et al., 2019 CARDOSO et al., 2020; ALVES et al., 2020; MA et al., 2020). However, the expression analysis of these genes is necessary for determining the protein content due to some *cry*, and *cyt* genes are cryptic, and incomplete promoters and transposon sequences that may interfere with the production of pesticidal proteins (DANKOCSIK et al., 1990; RANG et al., 2015).

The identification of insecticidal proteins expression is essential to explain the differences in toxicity between strains with identical gene profiles and to understand why production batches of a *Bt*-based insecticide vary in their toxicity characteristics (RANG et al., 2015; CABALLERO et al., 2020). Thus, combining the use of genomic and proteomic analyses has enabled the determination of pesticide protein genes content and their expression (GOMIS-CEBOLLA et al., 2018; BARAGAMAARACHCH et al., 2019; KHORRAMNEJAD et al., 2021).

In this sense, this work aimed to explore the insecticidal potential of new *Bt* strains isolated in the Tocantins to control lepidopteran and dipteran insects, explore genes related to

pathogenicity through genome sequence, and identify the proteins expressed in the spores-crystals.

1.1 Objectives

1.1.1 *General objective*

This work aims to study the entomopathogenic potentiality of *Bt* strains against dipteran and lepidopteran insects and to characterize the genome and proteome of these strains to detect genes and proteins related to their pathogenicity.

1.1.2 *Specific objectives*

- Perform bioassays using *Bt* strains for the control of dipteran and lepidopteran insects;
- Analyze the morphology of crystals produced by the isolates (all chapters);
- Perform the genome sequencing of *Bt* isolates (all chapters);
- Perform genomic assembly and annotation (all chapters);
- Characterize pesticide genes and virulence factors (all chapters);
- Screening for novel pesticidal genes;
- Perform phylogenetic analysis (all chapters);
- Purification of pesticidal proteins in the late sporulation phase (chapters 2 and 4);
- Identify peptide sequence in the spore-crystal mixture from LC-MS/MS (chapters 2 and 4);
- Combine genomic and proteomic data to identify expressed proteins (chapters 2 and 4).

2 LITERATURE REVIEW

2.1 *Bacillus thuringiensis*

Bacillus thuringiensis (*Bt*) is an aerobic or anaerobic facultative, Gram-positive, spore-forming bacterial that presents toxicity and specificity in the infection of insect hosts of different orders (e.g., Diptera, Lepidoptera, Coleoptera, Hymenoptera, Orthoptera, Homoptera and Phthiraptera), and other organisms such as nematodes, mites, and protozoa (TORRES-QUINTERO et al., 2018; MALOVICHKO et al., 2019; BARANEK et al., 2020; NAIR et al., 2020).

The pathogenicity of *Bt* is mainly due to protein crystals production, known as δ-endotoxins (Cry and Cyt), secreted at the sporulation stage, and can present different shapes (bipyramidal, cuboidal, rhomboid, spherical, or ovoid) (HADI et al., 2020). The target insects ingest δ-endotoxins that are solubilized by the pH of the midgut and then activated by proteases. Following, the active toxins binding to bind to receptors (Cry) or directly on the membrane (Cyt), and it results in pore formation in the midgut, cytolysis, and death of insects (RISKUWA-SHEHU et al., 2019). Besides *Bt* being the main bioactive used in the formulation of biological insecticides, *Bt* is the source of δ-endotoxin genes used in the development of transgenics plant crops (*Bt*-crops) (FERNÁNDEZ-CHAPA et al., 2019).

In addition to Cry and Cyt proteins, Vegetative insecticidal proteins (Vip), produced by vegetative cells of some strains, also contribute to the toxic activity of *Bt* (AL-FAR, 2020). The Vip proteins had mechanisms of action different from the δ-endotoxins proteins because binding to the other receptors type. Combining Vip and Cry proteins has increased toxicity through synergistic action and delayed resistance of insects to *Bt* proteins (BENGYELLA et al., 2018).

Other virulence factors play a relevant role in the infectious process of *Bt* and are involved in spore germination and cell multiplication, resulting in sepsis and host death (VILAS-BÔAS et al., 2012; RAJPUT et al., 2020). Virulence factors such as metalloproteases, chitinases, enterotoxins, hemolysins, phospholipases, and proteases may contribute to the bacteria invasion phase and increase the cytotoxic properties, and activity of *Bt* (UILLEMET et al., 2010; MALOVICHKO et al., 2019).

All these proprieties make *Bt* an important biological control agent against insect vectors of human diseases and insect pests of agriculture.

2.2 Insecticidal proteins toxic against dipteran and lepdopteran insects

Many *Bacillus thuringiensis* (*Bt*) subspecies have activity against insects of the Diptera order, which *include Bt*, *israelensis Bt jegathesan*, *Bt darmstadiensis*, *Bt kyushensis*, *Bt medellin*, *Bt fukuokaensis*, and *Bt higo*. These subspecies harbored different pesticidal proteins genes with activity against dipteran species, such as *cry1*, *cry2*, *cry4*, *cry10*, *cry11*, *cry19*, *cry20*, *cry24*, *cry27*, *cry30*, *cry39*, *cry44*, *cry47*, *cry50*, *cry54*, *cry56*, *mpp60*, *tpp80*, *cyt1*, and *cyt2* (VALTIERRA-DE-LUIS et al., 2020).

The *Bt* israelensis (*Bti*) is the most promising subspecies against disease vector mosquitoes, characterized by its unique protein content and high efficacy. *Bti* was the first subspecie reported with high insecticidal activity for larvae of mosquito vectors of several tropical diseases, such as *Aedes aegypti* and *Culex quinquefasciatus* (LOPES et al., 2019). Thus, *Bti* has been one of the best alternatives to control this type of insect, ensuring safety and prolonged control, especially in countries considered endemic (NAKAZAWA et al., 2020). However, it is necessary to carefully consider the risk analyses of bioinsecticides *Bti* based, since they can reach organisms that are not biological control targets, such as aquatic organisms (BORDALO et al., 2019).

Among the pesticidal proteins secreted by *Bti* are four Cry types (Cry4Aa, Cry4Ba, Cry10Aa, Cry11Aa) and three Cyt types (Cyt1Aa, Cyt2Ba, and Cyt1Ca) present in the plasmid called pBtoxis (BERRY et al., 2002; STEIN etal., 2006; RAJPUT et al., 2020). The mixture of these pesticidal proteins has higher insecticidal action than when tested individually due to synergistic interactions between Cry and Cyt proteins, especially Cyt1Aa (VALTIERRA-DE-LUIS et al., 2020). This synergism has explained by the difference between the action mechanisms of the Cry and Cyt proteins, increasing the *Bti* toxicity. While Cry binds to specific receptors in the intestinal cells of insects, Cyt forms pores and can have a deterrent effect directly on the midgut membrane (TETREAU et al., 2020; ONOFRE et al., 2020). The Cry10Aa and Cyt2Ba proteins demonstrated high activity against *A. aegypti* larvae, while Cry4Aa, Cry4Ba, or Cry11Aa interact synergistically with Cyt1Aa against the same insect species (VALTIERRA-DE-LUIS et al., 2020).

Mpp60A and Mpp60B proteins (formerly Cry60A/Cry60B) also have been detected in some *Bti* strains (CABALLERO et al., 2020). These proteins belong to Etx/Mtx2 protein family and have been found in other subspecies, such as *Bt jegathesan* and *Bt malayensis* (VALTIERRA-DE-LUIS et al., 2020).

In agriculture, *Bt* has been used as a source of pesticidal proteins to produce transgenic crops and as formulations for direct application to fields (AL-FAR, 2020). Among the types of pests that affect vegetable crops, lepidopteran insects have been the target of programs using *Bt* due to damages and economic losses in agriculture (AL-FAR, 2020). These insects' pests are susceptible to proteins such as *cry1*, *cry2*, *cry7B*, *cry8D*, *cry9*, *cry15A*, *cry19*, *cry20*, *cry22A*, *cry32A*, *cry51A*, *cry54*, and *cry59* (FERNÁNDEZ-CHAPA et al., 2019).

Bt kurstaki (*Btk*) represents the main subspecies with toxicity against lepidopteran (SMITHA et al., 2020; KIM and KIM, 2020), and most of the bioinsecticides using to lepidopteran-type pests are composed of *Btk* (NAWROT-ESPOSITO et al., 2020). The commercial strain *Btk* HD-1 produces Cry1Aa, Cry1Ab, Cry1Ac, Cry1Ia, Cry2Aa, and Cry2Ab proteins (IBRAHIM et al., 2010). Other *Btk* isolates have shown similar or variable content to the HD-1 strain. For example, *Btk* SA-11 produces six proteins (Cry1Aa, Cry1Ab, Cry1Ac, Cry1Ia, Cry2Aa and Cry2Ab), *Btk* Y*BT*-1520 procudes five pesticidal proteins (Cry1Aa, Cry1Ab, Cry1Ac, Cry2Aa and Cry2Ab), while the parasporal crystal of *Btk* HD-73 contains only the Cry1Ac protein (YU et al., 2020; QIU et al., 2010).

The Cry1Ac toxin is one of the most active against lepidopteran larvae among the Cry1 protein types. The *Chrysodeixis includens* and *Anticarsia gemmatalis* are the most economically relevant soybean pests in the American continent, and the transgenic soybean (Glycine max) expressing Cry1Ac protein is commercialized to control these insects (MUSHTAQ, SHAKOORI, JURAT-FUENTES, 2018).

The Cry2Ab and one or more Cry1 toxins have been combined in transgenic plants (FABRICK et al., 2020). Proteins of the Cry1I type, such as Cry1Ia, are known to have specificity for insects of the Coleoptera and Lepidoptera orders. This class of pesticidal proteins differs from other Cry proteins due are not found as crystal constituents in *Bt*, being secreted in the early stationary phase (BERRETTA et al., 2020). Cry1I proteins are interesting for resistance management programs since they do not share binding receptors in the insect midgut with Cry1Ab or Cry1Ac (BERRETTA et al., 2020).

In addition to pesticidal Cry proteins, *Btk* isolates can produce the Vip3 proteins (YU et al., 2020; QIU et al., 2010). The Vip and Cry proteins do not share receptors, and their mixture may have a synergistic activity (FIGUEIREDO et al., 2019; AL-FAR, 2020). The Cry1Ac, Cry2Aa, and Vip3Aa had a synergistic effect toward *Chrysodeixis includens*, while Vip3Aa and Cry2Aa exhibited synergistic interaction against *Anticarsia gemmatalis* (BENGYELLA et al., 2018).

Despite the success of the *Bt*-crops, insect resistance to pesticidal proteins has been reported. The mutations or changes in the expression of receptors of toxins in the midgut of insects, such as cadherin-like proteins, alkaline phosphatase (ALP), aminopeptidase N (APN), and ABC transporters, may mediate the insect resistance (LIU et al., 2020; JURAT-FUENTES; HECKEL; FERRÉ, 2021).

In the field, the insect's resistance has involved alterations in Cry binding proteins that result in cross-resistance to toxins sharing the same binding site (JURAT-FUENTES; HECKEL; FERRÉ, 2021). For example, the Cry1Ac expressed in cotton was the target for resistance in insects due to mutations in the cadherin of Platyedra gossypiella, and this alteration also may affect the activity of other Cry1A proteins in this insect (LIU et al., 2020; JURAT-FUENTES; HECKEL; FERRÉ, 2021). Products containing multiple *Bt* toxins belonging to the same family also have been targeted for insect resistance. The Plutella xylostella has evolved in resistance to *Bt* sprays based on *Btk* or *Bt aizawai* (*Bta*) under field conditions. The Cry1A proteins present in *Btk* and *Bta* have explained the cross-resistance between the formulations of this species (TABASHNIK, et al., 1993).

Concerning Vip proteins, the alleles for Vip3Aa resistance have been detected in insect species, but no cases of field resistance have been reported (JURAT-FUENTES; HECKEL; FERRÉ, 2021).

The MIR162, a maize that expresses Vip3Aa20 commercialized for more than ten years, is an example of the success of the Vip3 protein. This crop is highly efficacious against field populations of lepidopteran pests from different geographic locations (WEN et al., 2023). Vip3Aa has multiple cleavage activation sites in the loop region between domain I and domain II, and it might interact with variable protease compositions in the midgut of different insect hosts. Since the proteolytic activity is also involved in pest insect resistance to Cry and Vip proteins, the multiple cleavage sites of Vip3 may manage resistance in insect pests caused to reduced protease activity in the midgut (JIANG et al., 2023).

The finding of new cry genes can help minimize resistance by targeting insects to pesticidal proteins commercialized (FAYAD et al., 2019; ZHOU et al., 2020). Moreover, to potentiate the use of *Bt* pesticidal proteins, a mixture of those with little structural similarity and varied mechanisms of action is necessary. It ensures potent synergistic between proteins and allows control of insect resistance (AL-FAR, 2020).

2.3 Genomic of *Bacillus thuringiensis*

Since the advent of DNA sequencing technologies and their cost-effectiveness, genomes of *Bacillus thuringiensis* (*Bt*) strains have been sequenced. Several projects have focused on sequencing, assembly, and annotation of *Bt* genomes from different geographic areas (JOUZANI; VALIJANIANIAN; SHARAFI, 2017; MA et al., 2020; CARDOSO et al., 2020). The genomic characterization allowed the detection of the entire pathogenicity-related genes, including novel pesticidal proteins (DOMÍNGUEZ-ARRIZABALAGA et al., 2019).

In the study by Zhou et al. (2020), a novel pesticidal gene was found in a *Bt* strain and assigned to encode the Epp protein that showed toxicity against Spodoptera litura and Culex pipiens pallens. Bioinformatics analyses also showed that despite the low gene homology, Epp has levels of similarity in structural terms when compared to Cry protein structures. In another study, genome sequencing of the *Bt* isolate BLC406 revealed different patterns in its cry gene content, indicating the presence of five cry genes (*cry11, cry22, cry2, cry60, cry64*) and two *vip4* genes (ZGHAL et al., 2018).

Jeong, Choi, and Park (2017) reported the genome sequence of *Bt* kurstaki (*Btk*), BP865 an isolate with toxic activity against lepidopteran pests. In addition to identifying insecticidal genes (*cry1Aa*, *cry1Ab*, *cry1Ac*, *cry1Ia*, and *cry2Ab*), genome sequencing of *Btk* BP865 allowed the detection of two genes coding non-ribosomal peptide synthase (NRPS), which has pharmacological properties, which may extend its use beyond insect biological control.

Fayad et al. (2019), isolated a strain of Lebanese *Bti*, with activity insecticide tested against lepidopteran and dipteran larvae. The isolate was sequenced, and its genome was shown to possess a plasmid with an additional, functional gene encoding for Cry4Ba, which justifies its greater toxic potential in mosquito control with respect to a commercial strain. Another recent study involving the sequencing of the *Bti* LLP29 isolate made it possible not only to analyze pesticide genes but also genes involved in stress conditions such as UV, heat, hyperoxide, and high salinity. In addition, the *cry22Aa* gene, toxic to insects of Lepidoptera, Coleoptera, and Hymenoptera, but not to Diptera, was also detected in *Bti* LLP29 (MA et al., 2020).

In addition to the characterization of the pesticidal Cry and Cyt proteins, the genome sequence of *Bt* isolates has allowed a better understanding of the plasmid pattern. Sequencing of several representative *Bti* isolates has demonstrated that there may be little heterogeneity in the plasmid content and structure of the strains, as well as allowing the calculation of plasmid copy number and the identification of a novel 360 kb plasmid (BOLOTIN et al., 2017). The

sequencing of *Bti* strains isolated from Brazil and exhibiting toxicity against disease vector mosquitoes has enabled comparative analysis based on SNPs (Single Nucleotide Polymorphism). These results indicated high chromosomal conservation and plasmid content in this subspecie (ALVES et al., 2020).

Many *Bt* strains have cryptic plasmids, defined as small circular sequences without function defined. Genomic sequencing may be a tool to study plasmid cryptic, indicating its role and importance in the development and survival of *Bt* within its ecological niche. Cardoso et al. (2020), for example, analyzed the sequence of the plasmid pAW63 from the *Btk* isolate HD73 and characterized its Rap63-Phr63 system as a sporulation regulator of the isolate, highlighting its importance in the development and survival of *Bt* within its ecological niche (CARDOSO et al., 2020).

Besides plasmids, other types of transposable elements, such as ISs (Insertion Sequences) and Transposases, have been studied through genomic analysis. The genome of *Btk* isolate Y*BT*-1520 revealed a high number of ISs (Insertion Sequences) when compared to 18 genomes of the Bacillus cereus group, including different *Bt* strains (QIU et al., 2010). This study provides data that will help to understand IS-mediated genomic rearrangements in *Bt* and other B. cereus species. Already, genomic analysis of the *Bt* isolate Bc601 demonstrates that genes in its extrachromosomal content mainly encode transposases, transcriptional regulators, recombinases, type VII secretion proteins, and cell surface proteins (JIA et al., 2016).

2.4 Proteomic studies of *Bacillus thuringiensis*

A proteome is the total proteins located in a cell, tissue, organ, an entire individual, or the environment at a given time (SARETHY and SAHARAN, 2021). Since proteins represent functional molecules, proteomics-based studies allow studying the role of individual proteins in the life cycle of an organism (SARETHY and SAHARAN, 2021). Combining proteomics with other molecular techniques may be an effective tool to identify proteins with a biotechnological value that can be differentially expressed (RAM et al., 2018).

The identification of pesticidal proteins present in mixtures of spores and crystals of *Bacillus thuringiensis* (*Bt*) strains is still limited. The commercialized *Bt*-based formulations have no description of the composition of active ingredients, of which their characterization is based only on the insecticidal potential against a particular insect species. According to

Caballero et al. (2020), the insecticidal potential of *Bt* strains based only on bioassay is disadvantageous because its potency depends on the insect species tested.

Although genome analysis is a powerful tool for studying genes related to the pathogenicity of *Bt*, the analysis of the expression of these genes is an indispensable step for determining the protein content produced. Some *cry* and *cyt* genes are considered cryptic, i.e., not be expressed or have insignificant levels of expression (DANKOCSIK et al., 1990). Besides, incomplete promoters and transposon sequences may also interfere with the production of pesticidal proteins (RANG et al., 2015).

In this regard, the identification of insecticidal proteins expression is essential to infer the insecticidal activity of *Bt* strains, to explain the differences in toxicity between strains with identical gene profiles, and to understand why production batches of a *Bt*-based insecticide vary in their toxicity characteristics (CABALLERO et al., 2020). Thus, in association with genomics, proteomics has been applied in biological control studies involving *Bt* strains. In this regard, proteomic approaches have been used for studies of *Bt* isolates to investigate expression modulation and identify insecticidal proteins and other expressed virulence factors, such as their relative abundance.

The proteomics study, which investigated the impacts of long-term heat stress on physiological processes and biochemical contents of the *Bt* YBT-1520, revealed strategies for survival (WU et al., 2011). Heat stress mainly influenced the abilities to synthesize insecticidal proteins and other potential pathogenic factors, cell adhesion and motility, sporulation, and positive accumulation of poly (3-hydroxybutyrate) (PHB). Crystal proteins were below the detectable level, and the expression was decreased for pathogenic factors (NprB, CalY, InhA, SLP, and FlaA). The key survival strategy of *Bt* Y*BT*-1520 to long-term heat stress was to regulate BDH1, GuaB, and PepA enzymes, metabolic enzymes, to reduce metabolic load, and to increase PHB synthesis and accumulation (WU et al., 2011).

Combining genomic and proteomic analyses is an effective tool for determining pesticidal protein genes expressed and their relative abundance. Genomic analysis of the *Bt* 4.0718 detected the presence of five Cry proteins (*cry1Aa*, *cry1Ac*, *cry1Ia*, *cry2Aa*, *and cry2Ab*) (RANG et al., 2015). However, proteomic data from this *Bt* strain o*Bt*ained from proteins collected at vegetative and sporulation stages indicated the presence of only Cry1Aa, Cry1Ac, and Cry2Aa proteins (HUANG et al., 2012). Cry2Ab was not expressed due to an incomplete promoter, while Cry1Ia expression was affected by a transposon sequence (RANG et al., 2015).

The sequencing of the complete genome of *Bt* strain X022 and their comparison with proteomic data at the spore release stage indicated the presence of *cry1Ac*, *cry1Ia*, *cry2Ab*, and *vip3A* genes, and the expression of Cry1Ca, Cry1Ac, and Cry1Da (LIU et al., 2015; QUAN et al., 2016). The Cry1Ia and Vip3A could not be detected in the proteomic data due to their production at the vegetative phase. The presence of Cry1Ca and Cry1Da proteins only in proteomics data indicated assembly failures in plasmids, and it was confirmed using PCR (Polymerase Chain Reaction) (QUAN et al., 2016).

The genomics and proteomics analysis of *Bt* strains isolated from Spain revealed new pesticidal proteins and different expression levels. The isolate E-SE10.2 expressed two new vegetative proteins, Vip2Ac-like_1 and Sip1Aa-like_1, which showed no differences in expression at 24 h and 48 h, and the parasporal crystal produced contained only one protein, Cry23Aa-like. The isolate O-V84.2 encodes three new vegetative proteins, Vip4Aa-like_1, Vip4Aa-like_2, and Vip2Ac-like_2, in which the Vip4Aa-like_1 protein was twice the most abundant at 24 h than at 48 h. The Vip4Aa-like_2 was detected only at 24 h, and the expression of Vip2Ac-like_2 had no differences at 24 h and 48 h. The parasporal crystal of O-V84.2 showed three types of proteins: Cry45Aa-like, Cry32-like, and Cry73-like (GOMIS-CEBOLLA et al., 2018).

The study of the *Bt* AB1 strain, toxic against Plutella xilostela, showed that among the annotated insecticidal protein-coding genes of the AB1 genome (*cry1Aa*, *cry1Ca*, *cry1Da*, *cry1Ia*, *cry2Ab*, and *cry9*), the pesticidal proteins Cry1Ca and Cry1Da accounted for most of the toxin fraction in parasporal crystals (BARAGAMAARACHCH et al., 2019).

Genome sequencing of the *Bt* isolate KhF, toxic to Plodia interpunctella and Grapholita molesta larvae, revealed that the *Bt* KhF strain contained nine coding sequences with homologies to *Bt* insecticide genes. The mixture of spores and crystals from *Bt* KhF was subjected to liquid chromatography and tandem mass spectrometry (LC-MS/MS) to evaluate the protein composition. The results revealed two new pesticidal proteins, KhFB and KhFA, which have founded in abundance in the crystals of the *Bt* KhF strain. In addition, the proteomic analysis also indicated 327 different proteins, including virulence factors, which may also influence the pathogenicity of the KhF strain (KHORRAMNEJAD et al., 2021).

3 REFERENCES

BERRY, C., O'NEIL, S., BEN-DOV, E., JONES, A. F., MURPHY, L., QUAIL, M. A., PARKHILL, J. Complete sequence and organization of pBtoxis, the toxin-coding plasmid of *Bacillus thuringiensis* subsp. *israelensis*. **Applied and Environmental Microbiology**, v. 68, p. 5082-5095, 2002.

MUSHTAQ, R., SHAKOORI, A. R., JURAT-FUENTES, J. L. Domain III of Cry1Ac is critical to binding and toxicity against soybean looper (*Chrysodeixis includens*) but not to velvetbean caterpillar (*Anticarsia gemmatalis*). **Toxins**, v. 10, n. 3, p. 95, 2018.

CASTRO, B. M. D. C. E., MARTINEZ, L. C., BARBOSA, S. G., SERRÃO, J. E., WILCKEN, C. F., SOARES, M. A., ZANUNCIO, J. C. Toxicity and cytopathology mediated by *Bacillus thuringiensis* in the midgut of *Anticarsia gemmatalis* (Lepidoptera: Noctuidae). **Scientific reports**, v. 9, n. 1, 6667, 2019.

BENGYELLA, L., YEKWA, E. L., IFTIKHAR, S., NAWAZ, K., JOSE, R. C., FONMBOH, D. J., ROY, P. Global challenges faced by engineered *Bacillus thuringiensis* Cry genes in soybean (Glycine max L.) in the twenty-first century. 3 Biotech, v. 8, p. 1-15, 2018.

DANKOCSIK C., DONOVAN W. P., JANY C. S. Activation of a cryptic crystal protein gene of *Bacillus thuringiensis* subspecies *kurstaki* by gene fusion and determination of the crystal protein insecticidal specificity. **Molecular Microbiology,** v. 4, p. 2087–2094, 1990.

VALTIERRA-DE-LUIS D., VILLANUEVA M., BERRY C., CABALLERO P. Potential for *Bacillus thuringiensis* and Other Bacterial Toxins as Biological Control Agents to Combat Dipteran Pests of Medical and Agronomic Importance. **Toxins**, v.12, n. 12, p. 773, 2020.

CABALLERO, J., JIMÉNEZ-MORENO, N., ORERA, I., WILLIAMS, T., FERNÁNDEZ, A. B., VILLANUEVA, M., ANCÍN-AZPILICUETA, C. Unraveling the composition of insecticidal crystal proteins in *Bacillus thuringiensis*: A proteomics approach. **Applied and Environmental Microbiology**, v. 86, n. 12, p. e00476-20, 2020.

AL-FAR, I. M. *Bacillus thuringiensis* and its pesticidal crystal proteins. **International Journal of Fauna and Biological Studies**, n. 7, v. 3, p. 157-162, 2020.

ALVES, G. B.; MELO, F. L.; OLIVEIRA, E. E.; HADDI, K.; COSTA, L. T.; DIAS, M. L.; SANTOS, G. R. Comparative genomic analysis and mosquito larvicidal activity of four *Bacillus thuringiensis* subsp. *israelensis* strains. **Scientific reports**, v. 10, n. 1, p. 1-12, 2020.

BARAGAMAARACHCHI, R. Y., SAMARASEKERA, J. K. R. R., WEERASENA, O. V. D. S. J., LAMOUR, K., JURAT-FUENTES, J. L. Identification of a native *Bacillus thuringiensis* strain from Sri Lanka active against Dipel-resistant *Plutella xylostella*. **PeerJ**, v. 7, p. e7535, 2019.

BARANEK, J.; POGODZIŃSKI, B.; SZIPLUK, N.; ZIELEZINSKI, A. TOXiTAXi: a web resource for toxicity of *Bacillus thuringiensis* protein compositions towards species of various taxonomic groups. **Scientific Reports**, v. 10, n. 1, p. 1-12, 2020.

BERRETTA, M. F., PEDARROS, A. S., SAUKA, D. H., PÉREZ, M. P., ONCO, M. I., BENINTENDE, G. B. Susceptibility of agricultural pests of regional importance in South America to a *Bacillus thuringiensis* Cry1Ia protein. **Journal of Invertebrate Pathology**, p. 107354, 2020.

BOLOTIN, A., GILLIS, A., SANCHIS, V., NIELSEN-LEROUX, C., MAHILLON, J., LERECLUS, D., SOROKIN, A. Comparative genomics of extrachromosomal elements in *Bacillus thuringiensis* subsp. *israelensis*. **Research in microbiology**, v. 168, n. 4, p. 331-344, 2017.

BORDALO, M. D.; GRAVATO, C.; BELEZA, S.; CAMPOS, D.; LOPES, I.; PESTANA, J. L. T. Lethal and sublethal toxicity assessment of *Bacillus thuringiensis* var. *israelensis* and Beauveria bassiana based bioinsecticides to the aquatic insect Chironomus riparius. **Science of the Total Environment,** v. 698, p. 134155, 2020.

CARDOSO, P., FAZION, F., PERCHAT, S., BUISSON, C., VILAS-BÔAS, G., LERECLUS, D. Rap-Phr systems from plasmids pAW63 and pHT8-1 act together to regulate sporulation in the *Bacillus thuringiensis* serovar *kurstaki* HD73 strain. **Applied and Environmental Microbiology**, v. 86, n. 18, 2020.

DOMÍNGUEZ-ARRIZABALAGA, M.; VILLANUEVA, M.; FERNANDEZ, A. B.; CABALLERO, P. A Strain of *Bacillus thuringiensis* Containing a Novel cry7Aa2 Gene that Is Toxic to *Leptinotarsa decemlineata* (Say) (Coleoptera: Chrysomelidae). **Insects**, v. 10, n. 9, p. 259, 2019.

EVANGELISTA, J. G. V. N.; DOS SANTOS ANDRADE, A. P.; SANTOS, A. M.; SILVA, F. D. C. G.; CARMO, A. S.; ALVES, G. A. R.; ... DA SILVA BORGES, L. Inseticida biológico no controle de lagarta *Spodoptera eridania* no cultivo da soja. **Revista Ibero-Americana de Ciências Ambientais**, v. 12, n. 2, 2020.

FABRICK, J. A., LEROY, D. M., UNNITHAN, G. C., YELICH, A. J., CARRIÈRE, Y., LI, X., TABASHNIK, B. E. Shared and independent Genetic Basis of Resistance to *Bt* toxin Cry2Ab in Two Strains of Pink Bollworm. **Scientific reports**, v. 10, n. 1, p. 1-11, 2020.

FAYAD, N.; PATIÑO-NAVARRETE, R.; KAMBRIS, Z.; ANTOUN, M.; OSTA, M.; CHOPINEAU, J. AWAD, M. K. Characterization and Whole Genome Sequencing of AR23, a Highly Toxic *Bacillus thuringiensis* Strain Isolated from Lebanese Soil. **Current Microbiology**, v. 76, n. 12, p. 1503-1511, 2019.

FERNÁNDEZ-CHAPA, D.; RAMÍREZ-VILLALOBOS, J.; GALÁN-WONG, L. Toxic Potential of *Bacillus thuringiensis*: An Overview. In: **Protecting Rice Grains in the Post-Genomic Era**. IntechOpen, p. 23, 2019.

FIGUEIREDO, C. S.; LEMES, A. R. N.; SEBASTIÃO, I.; DESIDÉRIO, J. A. Synergism of the *Bacillus thuringiensis* Cry1, Cry2, and Vip3 proteins in *Spodoptera frugiperda* control. **Applied biochemistry and biotechnology**, v. 188, n. 3, p. 798-809, 2019.

GOMIS-CEBOLLA, J., SCARAMAL RICIETTO, A. P., FERRÉ, J. A genomic and proteomic approach to identify and quantify the expressed *Bacillus thuringiensis* proteins in the supernatant and parasporal crystal. **Toxins**, v. 10, n. 5, p. 193, 2018.

HADI, H.W. Significance of plasmids in *Bacillus cereus* sensu lato group: a review. **Plant Archives**, v. 20, n. 1, p. 881-890, 2020.

HUANG, S., DING, X., SUN, Y., YANG, Q., XIAO, X., CAO, Z., XIA, L. Proteomic analysis of *Bacillus thuringiensis* at different growth phases by using an automated online two-dimensional liquid chromatography-tandem mass spectrometry strategy. **Applied and environmental microbiology**, v. 78, n. 15, p. 5270-5279, 2012.

IBRAHIM, M. A.; GRIKO, N.; JUNKER, M.; BULLA, L. A. *Bacillus thuringiensis*: a genomics and proteomics perspective. **Bioengineered bugs**, v. 1, n. 1, p. 31-50, 2010.

JEONG, H.; CHOI, S-K.; PARK, S-H. Genome sequences of *Bacillus thuringiensis* Serovar *kurstaki* strain BP865 and *B. thuringiensis* Serovar *aizawai* Strain HD-133. **Genome Announcements**, v. 5, n. 5, 2017.

JIA, N., DING, M. Z., GAO, F., YUAN, Y. J. Comparative genomics analysis of the companion mechanisms of *Bacillus thuringiensis* Bc601 and *Bacillus endophyticus* Hbe603 in bacterial consortium. **Scientific reports**, v. 6, n. 1, p. 1-8, 2016.

JIANG, K., CHEN, Z., SHI, Y., ZANG, Y., SHANG, C., HUANG, X., GAO, X. A strategy to enhance the insecticidal potency of Vip3Aa by introducing additional cleavage sites to increase its proteolytic activation efficiency. **Engineering Microbiology**, 100083, 2023.

JOUZANI, G. S.; VALIJANIAN, E.; SHARAFI, R. *Bacillus thuringiensis*: a successful insecticide with new environmental features and tidings. **Applied microbiology and biotechnology**, v. 101, n. 7, p. 2691-2711, 2017.

JURAT-FUENTES, Juan Luis; HECKEL, David G.; FERRÉ, Juan. Mechanisms of resistance to insecticidal proteins from *Bacillus thuringiensis*. **Annual Review of Entomology**, v. 66, p. 121-140, 2021.

KAYAM, D.; TIRUPATI, M. K.; KARANAM, H. P. Diversity of *Bacillus thuringiensis cry* genes in soils of Andhra Pradesh, India. **Indian Journal of Biochemistry and Biophysics (IJBB)**, v. 57, n. 4, p. 471-480, 2020.

KHORRAMNEJAD, A., GOMIS-CEBOLLA, J., TALAEI-HASSANLOUEI, R., BEL, Y., ESCRICHE, B. Genomics and proteomics analyses revealed novel candidate pesticidal proteins in a lepidopteran-toxic *Bacillus thuringiensis* strain. **Toxins**, v. 12, n. 11, p. 673, 2020.

KIM, H.; KIM, Y. Immunosuppressive Activity of an Entomopathogenic Bacteria, *Xenorhabdus ehlersii* KSY, and Its Application to Enhance Insecticidal Activity of *Bacillus thuringiensis*. **Korean journal of applied entomology**, v. 58, n. 2, p. 101-109, 2019.

TABASHNIK B. E.; FINSON N.; JOHNSON M. W; MOAR W. J. 1993. Resistance to toxins from *Bacillus thuringiensis* subsp. *kurstaki* causes minimal cross-resistance to *B. thuringiensis* subsp. *aizawai* in the diamondback moth (Lepidoptera: Plutellidae). **Applied and Environmental Microbiology**, v. 59, n. 5, p. 1332-1335, 1993.

LIU, Z., FU, S., MA, X., BAXTER, S. W., VASSEUR, L., XIONG, L., YOU, M.. Resistance to *Bacillus thuringiensis* Cry1Ac toxin requires mutations in two Plutella xylostella ATP-binding cassette transporter paralogs. **PLoS pathogens**, v. 16, n. 8, p. e1008697, 2020.

LOPES, R. P.; LIMA, J. B. P.; MARTINS, A. J. Insecticide resistance in *Culex quinquefasciatus* Say, 1823 in Brazil: a review. **Parasites & Vectors**, v. 12, n. 1, p. 591, 2019.

MA, W.; CHEN, H.; JIANG, X.; WANG, J.; GELBIČ, I.; GUAN, X.; ZHANG, L. Whole genome sequence analysis of the mosquitocidal *Bacillus thuringiensis* LLP29. **Archives of Microbiology**, p. 1-8, 2020.

MALOVICHKO, Y. V.; NIZHNIKOV, A. A.; ANTONETS, K. S. Repertoire of the *Bacillus thuringiensis* virulence factors unrelated to major classes of protein toxins and its role in specificity of host-pathogen interactions. **Toxins**, v. 11, n. 6, p. 347, 2019.

MEJIAS, L.; ESTRADA, M.; BARRENA, R.; GEA, T. A novel two-stage aeration strategy for *Bacillus thuringiensis* biopesticide production from biowaste digestate through solid-state fermentation. **Biochemical Engineering Journal**, p. 107644, 2020.

NAIR, K.; AL-THANI, R.; GINIBRE, C.; CHANDRE, F.; ALSAFRAN, M.; JAOUA, S. *Bacillus thuringiensis* strains isolated from Qatari soil, synthesizing δ-endotoxins highly active against the disease vector insect *Aedes aegypti* Bora Bora. **Heliyon**, v. 6, n. 10, p. e05003, 2020.

NAKAZAWA, M. M.; ARAÚJO, A. P.; MELO-SANTOS, M. A. V.; OLIVEIRA, C. M. F.; SILVA-FILHA, M. H. N. L. Efficacy and persistence of *Bacillus thuringiensis* svar. *israelensi*s (*Bti*) and pyriproxyfen-based products in artificial breeding sites colonized with susceptible or *Bti*-exposed *Aedes aegypti* larvae. **Biological Control**, v. 151, p. 104400, 2020.

NAWROT-ESPOSITO, M. P.; BABIN, A.; PASCO, M.; POIRIÉ, M.; GATTI, J. L.; GALLET, A. *Bacillus thuringiensis* bioinsecticides induce developmental defects in non-target *Drosophila melanogaster* larvae. **Insects**, v. 11, n. 10, p. 697, 2020.

POHARE, M. B.; WAGH, S. G.; UDAYASURIYAN, V. *Bacillus thuringiensis* as Potential Biocontrol Agent for Sustainable Agriculture. In: Current Trends in Microbial Biotechnology for Sustainable Agriculture. **Springer**, Singapore, p. 439-468, 2020.

QIU, N.; HE, J.; WANG, Y.; CHENG, G.; LI, M.; SUN, M.; YU, Z. Prevalence and diversity of insertion sequences in the genome of *Bacillus thuringiensis* Y*BT*-1520 and comparison with other *Bacillus cereus* group members. **FEMS microbiology letters**, v. 310, n. 1, p. 9-16, 2010.

QUAN, M., XIE, J., LIU, X., LI, Y., RANG, J., ZHANG, T., DING, X. Comparative analysis of genomics and proteomics in the new isolated *Bacillus thuringiensis* X022 revealed the

metabolic regulation mechanism of carbon flux following Cu^{2+} treatment. **Frontiers in Microbiology**, v. 7, p. 792, 2016.

RAJPUT, V. S.; JHALA, J.; ACHARYA, V. S. Biopesticides and their mode of action against insect pests: A review. **IJCS**, v. 8, n. 2, p. 2856-2862, 2020.

RAM, R. M., KESWANI, C., BISEN, K., TRIPATHI, R., SINGH, S. P., SINGH, H. B. Biocontrol technology: eco-friendly approaches for sustainable agriculture. In: Omics technologies and bio-engineering. **Academic Press**, 2018. p. 177-190.

RANG, J., HE, H., WANG, T., DING, X., ZUO, M., QUAN, M., XIA, L. Comparative analysis of genomics and proteomics in *Bacillus thuringiensis* 4.0718. **PLoS One**, v. 10, n. 3, p. e0119065, 2015.

RISKUWA-SHEHU, M. L.; NATA'ALA, M. K.; BABA, E. E. Biocontrol potential of *Bacillus thuringiensis* isolated from soil against mosquito larva. **South Asian Journal of Parasitology**, p. 1-7, 2019.

SARETHY, I. P.; SAHARAN, Anurag. Genomics, proteomics and transcriptomics in the biological control of plant pathogens: A review. **Indian Phytopathology**, v. 74, n. 1, p. 3-12, 2021.

SMITHA, R. B.; BENJAMIN, S.; PRAKASHKUMAR, R. Fermentation Strategies for the Production of α-Amylase and δ-endotoxin from *Bacillus thuringiensis* subsp. *kurstaki*. **Avid Science Monograph Series**, 192 p., 2019.

STEIN, C. JONES, G. W., CHALMERS, T., BERRY, C. Transcriptional analysis of the toxin-coding plasmid pBtoxis from *Bacillus thuringiensis* subsp. *israelensis*. Applied and environmental microbiology, v. 72, n. 3, p. 1771-1776, 2006.

TORRES-QUINTERO, M. C.; GÓMEZ, I.; PACHECO, S.; SÁNCHEZ, J.; FLORES, H.; OSUNA, J.; BRAVO, A. Engineering *Bacillus thuringiensis* Cyt1Aa toxin specificity from dipteran to lepidopteran toxicity. **Scientific reports**, v. 8, n. 1, p. 1-12, 2018.

UILLEMET, E., CADOT, C., TRAN, S. L., GUINEBRETIÈRE, M. H., LERECLUS, D., RAMARAO, N. The InhA metalloproteases of *Bacillus cereus* contribute concomitantly to virulence. **Journal of bacteriology**, v. 192, n. 1, p. 286-294, 2010.

VALTIERRA-DE-LUIS, D.; VILLANUEVA, M.; LAI, L.; WILLIAMS, T.; CABALLERO, P. Potential of Cry10Aa and Cyt2Ba, Two Minority δ-endotoxins Produced by *Bacillus thuringiensis* ser. *israelensis*, for the Control of *Aedes aegypti* Larvae. **Toxins**, v. 12, n. 6, p. 355, 2020.

VILAS-BÔAS, G. T.; ALVAREZ, R. C.; DOS SANTOS, C. A.; VILAS-BOAS, L. A. Fatores de virulência de *Bacillus thuringiensis*: o que existe além das proteínas Cry. **EntomoBrasilis**, v. 5, n. 1, p. 1-10, 2012.

WEN, Z., CONVILLE, J., MATTHEWS, P., HOOTMAN, T., HIMES, J., WONG, S., BRAMLETT, M. More than 10 years after commercialization, Vip3A-expressing MIR162 remains highly efficacious in controlling major Lepidopteran maize pests: laboratory

resistance selection versus field reality. **Pesticide Biochemistry and Physiology**, 105385, 2023.

WU, D., HE, J., GONG, Y., CHEN, D., ZHU, X., QIU, N., YU, Z Proteomic analysis reveals the strategies of *Bacillus thuringiensis* YBT-1520 for survival under long-term heat stress. **Proteomics**, v. 11, n. 13, p. 2580-2591, 2011.

YU, H.; YANG, C. J.; LI, N.; ZHAO, Y.; CHEN, Z. M.; YI, S. J.; HUANG, G. H. Novel strategies for the biocontrol of noctuid pests (Lepidoptera) based on improving ascovirus infectivity using *Bacillus thuringiensis*. **Insect Science**, 2020.

ZGHAL, R. Z.; GHEDIRA, K.; ELLEUCH, J.; KHARRAT, M.; TOUNSI, S. Genome sequence analysis of a novel *Bacillus thuringiensis* strain BLB406 active against *Aedes aegypti* larvae, a novel potential bioinsecticide. **International Journal of Biological Macromolecules**, v. 116, p. 1153-1162, 2018.

ZHOU, Y.; WU, Z.; ZHANG, J.; WAN, Y.; JIN, W.; LI, Y.; FANG, X. *Bacillus thuringiensis* novel toxin Epp is toxic to mosquitoes and *Prodenia litura* larvae. **Brazilian Journal of Microbiology**, p. 1-9, 2020.

CHAPTER I: Comparative genomic analysis and mosquito larvicidal activity of four *Bacillus thuringiensis* serovar *israelensis* strains

Published in: Scientific Report. https://doi.org/10.1038/s41598-020-60670-7 (Annex A)

Abstract

Bacillus thuringiensis serovar *israelensis* (*Bti*) is used to control insect vectors of human and animal diseases. in the present study, the toxicity of four strains of *Bti*, named T0124, T0131, T0137, and T0139, toward *Aedes aegypti* and *Culex quinquefasciatus* larvae was analyzed. The T0131 strain showed the highest larvicidal activity against *A. aegypti* (LC_{50}= 0.015 µg/ml) and *C. quinquefasciatus* larvae (LC_{50}= 0.035 µg/ml) when compared to the other strains. Furthermore, the genomic sequences of the four strains were obtained and compared. These *Bti* strains had chromosomes sizes of approximately 5.4 Mb with GC contents of ~35% and 5472–5477 putative coding regions. Three small plasmids (5.4, 6.8, and 7.6 kb) and three large plasmids (127, 235, and 359 kb) were found in the extrachromosomal content of all four strains. the SNP-based phylogeny revealed close relationship among isolates from this study and other *Bti* isolates, and SNPs analysis of the plasmids 127 kb did not reveal any mutations in δ-endotoxins genes. This newly acquired sequence data for these *Bti* strains may be useful in the search for novel insecticidal toxins to improve existing ones or develop new strategies for the biological control of important insect vectors of human and animal diseases.

1 Introduction

During sporulation, the gram-positive bacterium *Bacillus thuringiensis* (*Bt*) produces crystalline inclusions consisting of δ-endotoxins (Cry or Cyt proteins) with insecticidal activity[1]. Genomic analysis has contributed to the identification of new genes coding for toxins that are active against different insect species including orders such as Lepidoptera, Diptera[2–7], and Coleoptera[8]. Proteins with nematicida[l9–11] and molluscicidal[12] activities have also been described. In addition, genome sequencing of *Bt* strains with diverse ecological functions has been conducted, including a endophytic strain with potential utility in the biocontrol of phytopathogens[13].

Sequencing of complete *Bt* genomes has allowed structural and functional analysis of new plasmids that enhance our knowledge of the pathogenic properties of *Bt* in targeting organisms[14–17]. One study reported the plasmid sequence of a *Bacillus thuringiensis* serovar *israelensis* (*Bti*) strain and revealed that it may produce up to seven crystal-forming toxins, named Cry4A, Cry4B, Cry10A, Cry11A, Cyt1A, Cyt2Ba, and Cyt1Ca, which are all encoded by genes found in a single 127923 bp plasmid called pBtoxis[18]. The average size of the complete genome sequences of *Bti* is 6.1 Mb, with ~35% GC content of the chromosomal DNA and an average of 6132 coding sequences[19,20]. Genome sequences of seven *Bti* isolates have been reported so far[19,21–24].

In this study, we sequenced the genomes of four *Bti* strains, specifically T0124, T0131, T0137, and T0139 that were collected from the soil of the Tocantins state in Brazil and determined their larvicidal activity against larvae of two important mosquito species of *A. aegypti* and *C. quinquefasciatus*. Then, to better characterize these strains, we performed comparative and phylogenetic analyses among their different genomes and compared the potential insecticidal toxin genes and other virulence factors of the four *Bti* strains with the commercial *Bti* strain H14. In case we identify high anti-mosquito activity with these strains, we believe these new data are useful in the continuous search for new insecticidal toxins to improve the existing ones or develop new strategies for the biological control of important insect vectors of human and animal diseases.

2 Methods

2.1 Isolation of *Bti* strains

B. thuringiensis serovar *israelensis* (serotypes H14) T0124, T0131, T0137, and T0139 strains were isolated from a soil sample collected in Tocantins state (Brazil) according to the previously described protocol40. The bacterial strains were cultured at 28 °C for 12 h using the streak plate method on Luria-Bertani (LB) solid medium (10 gL−1 tryptone, 5 gL−1 yeast extract, 10 gL−1 NaCl, and 20 gL−1 Agar). Single bacterial colonies of each strain were inoculated in the LB liquid medium at 28 °C with shaking for 16 h.

2.2 Spore-crystal protein preparation and crystal analysis by SDS–PAGE

Spore-crystal mixtures were obtained according to the protocol described previously[41]. For SDS-PAGE analysis, the crystals were purified using hexane and low speed centrifugation according to the previously described method[42]. Proteins were sus-pended in a small volume of phosphate-buffered saline (136 mM NaCl, 1.4 mM KH2PO4, 2.6 mM KCl, 8 mM Na2HPO4, and 4.2 ml H2O; pH 7.4), and fractionated by electrophoresis on 12% SDS-PAGE gels[43].

2.3 Scanning electron microscopy

The characterization of the spores and Cry proteins from the T0124, T0131, T0137, and T0139 strains was performed by scanning electron microscopy. The strains were cultivated in NYSM agar medium at 30 °C for 72 h, then a loop of the isolate was collected and diluted in sterile water. A volume of 100 µL of this dilution was deposited over metallic supports and dried for 24 h at 37 °C, covered with gold for 180 s using an Emitech apparatus (model K550; Quorum Technologies, Lewes, UK), and observed under a Zeiss scanning electron microscope (model DSM 962; Carl Zeiss AG, Oberkochen, Germany) at 10 or 20 Kv.

2.4 Mosquitoes and toxicity bioassays

The colonies of *A. aegypti* and *C. quinquefasciatus* were established from insects collected from the field in regions of transition between the urban and rural areas in the state of Tocantins, Brazil, (11°40′55.7″ latitude S, 49°04′3.9″ longitude W), where no insecticides have been used for the control of mosquitoes. The insects were maintained in the Entomology Laboratory of the Federal University of Tocantins, Gurupi Campus, according to the methodology described previously[44]. Adult mosquitoes were main-tained on a 10% aqueous sucrose solution and the blood of live Wistar rats (*Rattus norvegicus albinus*). The larvae were

reared in plastic containers (35 cm × 5 cm) and were fed a sterilized diet (80/20 mix of chick chow powder/yeast). All bioassays were conducted at 26 ± 1 °C, 60.0 ± 5% RH, with a 12 h light-dark photoperiod. All applicable international, national, and institutional guidelines for the care and use of animals were followed. Bioassays were conducted using the suspension isolated from the spore-crystal mixtures against third instar *A. aegypti* and *C. quinquefasciatus larvae*. The concentrations were determined as described previously[45]. Seven concentrations were used for each spore-crystal mixture from each strain. Sterile distilled water was used as a control, and the larval mortality was recorded 24 h post inoculation. Three replicates with 25 larvae for each spore-crystal mixture were performed for all tested concentrations and for the non-treated control group. The spore-crystal mixture from the H14 strain was used as a reference.

2.5 Whole genome sequencing, assembly, and annotation

Total genomic DNA was extracted and purified using a Wizard® Genomic DNA Purification Kit (Promega, Madison, WI, USA) according to the manufacturer's instructions. DNA concentration and purity were measured using a NanoDrop™ 8000 (Thermo Fisher Scientific, Waltham, MA, USA). Whole genome sequencing was performed on the Illumina MiSeq™ platform using a paired-end application (2 × 150 bp) (Illumina, San Diego, CA, USA). The read quality of the sequenced libraries was analyzed using FastQC software v 0.11.346 and sequence reads were trimmed to yield a minimum Phred quality score > 20. The genome assembly was performed using Geneious v 10.1.3[47]. The *de novo* assembly was performance using Geneious assembler with medium sensitivity settings and allowing contigs with matching ends to circularize. The linear contigs were extended. For this, the reads were mapped back to the linear contigs and the resulting contigs were used as seeds for another attempted assembly until no further extension. Finally, the extended linear contigs were aligned and reordered using as reference the genome *Bti* HD-789 (accession number CP003763) from the "map to reference" tool with minimum overlap identity of 85%. The circular contigs were used to investigate plasmid-like sequences, by matching them against plasmid bank with custom BLAST tool. Genome annotation was added by the NCBI Prokaryotic Annotation.

2.6 Comparative genomic and phylogenetic analysis

Comparative chromosome-sequences analysis among T0124, T0131, T0137, T0139 and reference HD-789 was performed using BRIG (BLAST Ring Image Generator)[48]. The comparative analysis of the gene sequence of sporulation for the strains considered in this study and other species form the *Bacilli* group was performed using Blastx, and the heatmaps were generated using the MeV tool version 4.9[49]. Venn diagrams generation and orthologous cluster annotation for biological process, molecular function, and cellular component categories were achieved using OrthoVenn[50]. The collinearity and phylogenetic analysis among T0124, T0131, T0137, and T0139 and others 14 chromosomes of *Bt* was performed. The collinear analysis and display of results were conducted using Mauve with the parameters reported previously[51]. The phylogenetic tree based on single nucleotide polymorphisms (SNPs) was performed by CSI phylogeny 1.4 web[52] using the default parameters and HD-789 as reference. The SNPs were named, concatenated and aligned, and the tree was constructed using the maximum likelihood method. The phylogeny tree inferred was reviewed using MEGA X software[53] with 1000 replicates. The pBtoxis (NC_010076) was used as reference for the SNPs analysis of pT0124-4, pT0131-4, pT0137-4, and pT0139-4 using Geneious v 10.1.347, “Find SNPs/InDels” tool, with minimum coverage of 10, minimum variance frequency 0.75.

2.7 Nucleotide sequence accession number

The Whole Genome Shotgun projects of PRJNA521267, PRJNA521275, PRJNA521276, and PRJNA521307 *Bti* strains were deposited in DDBJ/ENA/GenBank under the accession numbers CP037890, CP035735, CP035736, and CP035737.

2.8 Statistical analyses

Concentration–mortality curves were estimated via probit analysis using the PROBIT procedure in the SAS statistical software package[54]. The differential susceptibility among mosquito species to H14 and the T0124, T0131, T0137, and T0139 *Bti* strains was assessed based on the estimated LC_{50} (i.e., the lethal concentration capable of killing 50% of tested mosquito species) of each strain and the toxicity ratios (TR_{50}) were estimated by dividing the LC_{50} value obtained for the T0124, T0131, T0137, and T0139 *Bti* strains by the LC_{50} value obtained for the H14 standard strain[55]. The 95% confidence intervals estimated for these toxicity rates were considered to be significantly different if they did not include the value 1[55].

3 Results

3.1 Larvicidal activity and features of δ-endotoxins

The spore-crystal mixtures of *Bti* strains T0124, T0131, T0137, and T0139 were tested against third instar larvae of *A. aegypti* and *C. quinquefasciatus*. The T0131 strain showed the highest larvicidal activity against *A. aegypti* (LC_{50}= 0.015 µg/ml) and *C. quinquefasciatus* (0.035 µg/ml) when compared to the other strains. Moreover, based on the toxicity ratios, the T0131 strain pre-sented similar toxicity to the reference strain H14 (Toxicity ratio = 1.1 against *A. aegypti* and Toxicity ratio = 1.3 against *C. quinquefasciatus*) (Table 1). The T0124, T0137, and T0139 isolates showed lower toxicities compared to the H14 strain (Table 1).

Table 1. Lethal concentration of *Bti* strains to larvae of *A. aegypti* and *C. quinquefasciatus*, isolated in the town of Gurupi-TO, Brazil.

Insecticide type	Strains	No. of insects	LC_{50} (95% FI[a]) µg /ml	LC_{95} (95% FI[a]) µg /ml	TRb_{50} (95% CL)	X^2	*P*
A. aegypti	T0124	525	0.069 (0.061 - 0.077)	0.243 (0.21 - 0.31)	5.2 (4.9 - 5.7)	4.36	0.36
	T0131	525	0.015 (0.012 - 0.018)	0.045 (0.03 - 0.07)	1.1 (0.9 - 1.3)	8.06	0.09
	T0137	525	0.165 (0.149 - 0.182)	0.534 (0.45 - 0.68)	12.6 (11.2 - 13.6)	5.06	0.28
	T0139	525	0.123 (0.096 - 0.157)	0.404 (0.28 - 0.79)	9.4 (8.2 - 10.3)	9.4	0.05
	H14	175	0.013 (0.011 - 0.016)	0.037 (0.03 - 0.05)	*	4.33	0.36
C. quinquefasciatus	T0124	450	0.172 (0.157 - 0.188)	0.467 (0.39 - 0.59)	6.4 (5.7 - 6.8)	1.97	0.58
	T0131	450	0.035 (0.031 - 0.039)	0.101 (0.08 - 0.14)	1.3 (1.2 - 1.4)	0.72	0.86
	T0137	525	0.239 (0.219 - 0.261)	0.630 (0.54 - 0.78)	8.6 (7.9 - 9.4)	7.07	0.13
	T0139	375	0.250 (0.220 - 0.283)	0.791 (0.63 - 1.10)	9.3 (8.2 - 9.8)	3.36	0.19
	H14	175	0.028 (0.024 - 0.032)	0.069 (0.06 - 0.10)	*	6.49	0.16

[a]FI = Fiducial Intervals; [b]TR_{50} = Toxicity ratio determined by LC_{50} of given strain/LC_{50} of the reference strain H14 (*); 95% CL= 95% Confidence limits; χ^2 = *Chi*-square for lack-of-fit to the probit model, and *P* = Probability associated with the *chi*-square statistic.

However, the SDS-PAGE analysis of crystal protein content revealed that all the strains have similar protein profiles. δ-endotoxins with molecular weights of 130, 70, and 27 kDa (Fig. 1A) and round morphology (Fig. 1B–E), characteristics of the *Bti* protein profile, were observed for the strains.

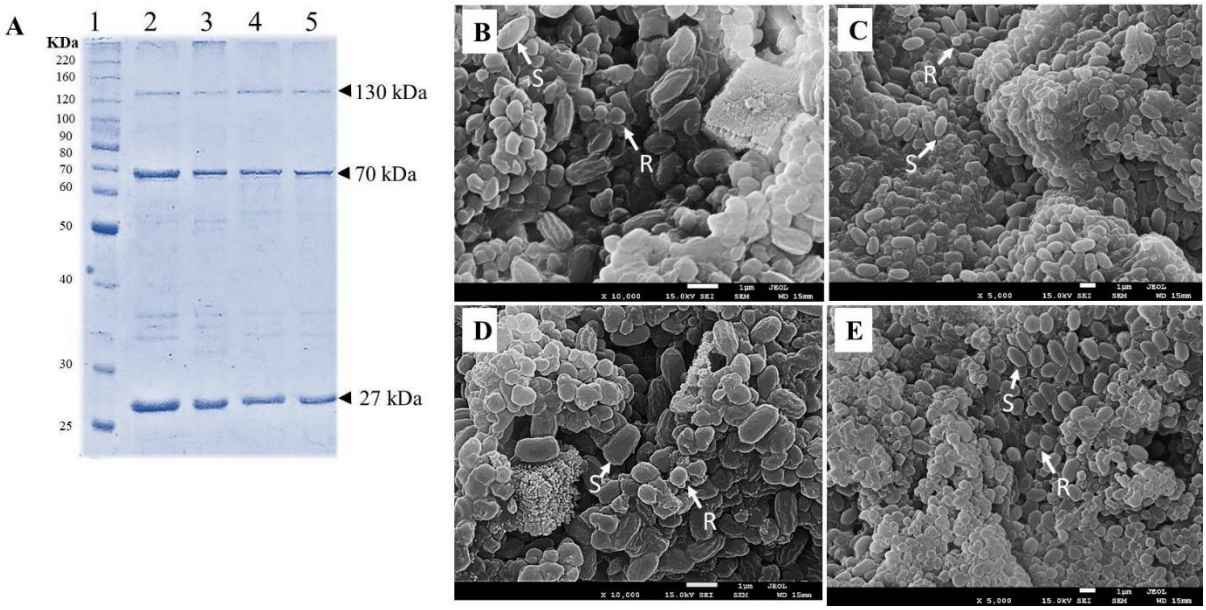

Figure 1. Crystal protein profile of *Bti* strains. Lane 1: molecular weight marker (Invitrogen); Lane 2: T0124; Lane 3: T0131; Lane 4: T0137; and Lane 5: T0139. Seven µg of solubilized crystals from each strain were analyzed by SDS-PAGE. The ultrastructural characterization of the spores and Cry proteins from T0124 (**B**), T0131 (**C**), T0137 (**D**), and T0139 (**E**) strains. All strains presented round crystals. Arrows indicate spores (S) and round crystals (R).

3.2 Genome features

The average size of the chromosomal draft sequences of the T0124, T0131, T0137, and T0139 isolates was 5.4 Mb, with GC contents of ~35%. Chromosomes of these isolates contained 5477 (T0124), 5473 (T0131) and 5472 (T0137 and T0139) protein-coding genes. The number of tRNA genes was consistent over the strains (122) while small variation was seen in the number of rRNA (39–42) among them (Table 2).

Table 2. General features of the genome sequences of *Bti* T0124, T0131, T0137 e T0139 strains.

General features	**T0124**	**T0131**	**T0137**	**T0139**
Average coverage (nº reads)	10.2	10.3	17.5	11.0
Chromosome size (bp)	5 415 530	5 414 369	5 414 369	5 414 367
Sites no cover (%)	0.7	0.1	0.06	0.1
GC content (%)	35.2	35.3	35.3	35.3
CDS	5477	5473	5472	5472
tRNA	112	112	112	112
rRNA	39	42	42	42
Plasmids (nº)	6	6	6	6
Chromossome	1	1	1	1

Genome assembly showed high similarities in nucleotide identity with the HD-789 reference genome (Fig. 2). Regions (in blank) within positions 3 407 568 – 3 451 845 bp and 4 278 610 – 4 319 513 bp indicate two prophage sequences located on the reference chromosome which are absent in the T0124, T0131, T0137, and T0139 isolates (Fig. 2).

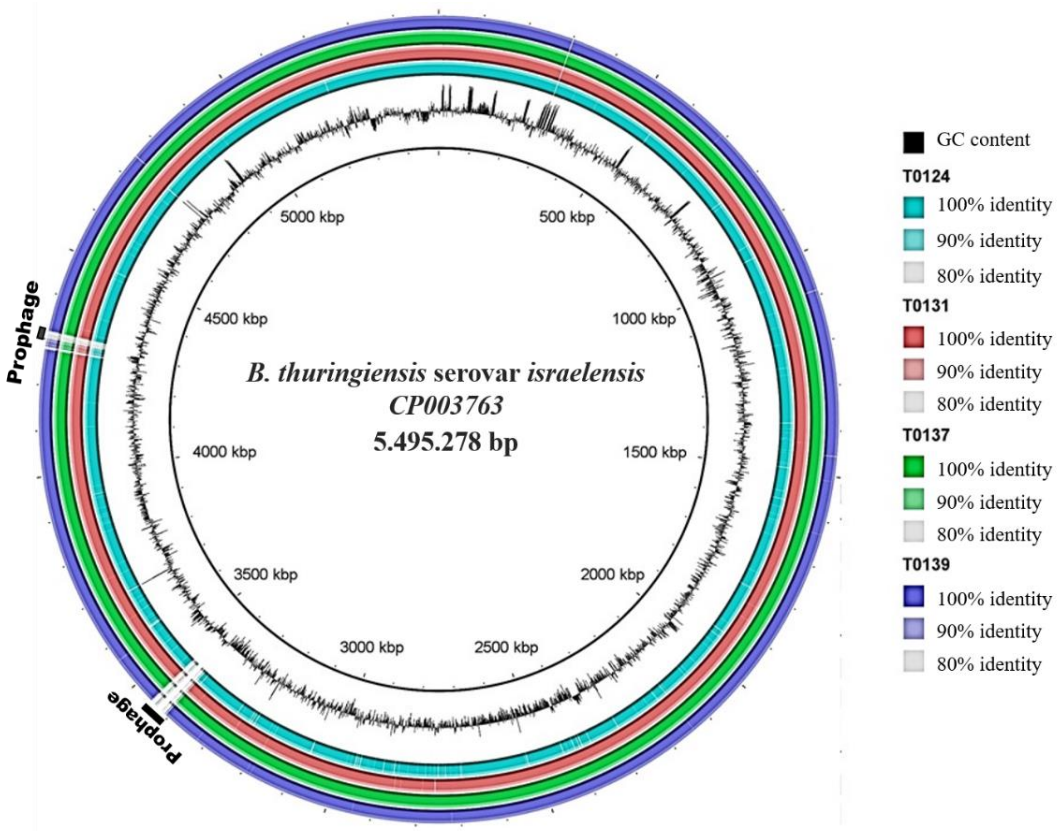

Figure 2. Comparative chromosomal nucleotide analysis of *Bti* strains. The concentric rings represent the sequences of T0124, T0131, T0137, and T0139 against the reference CP003763 strain. The black circle represents GC content of CP003763. The blue circle represents T0124, red circle represents T0131, green circle represents T0137, and purple circle represents T0139. Regions with less than 80% identity appear as blanks on each ring. This circular map was generated using the BLAST Ring Image Generator (BRIG) tool[48].

All four strains contain six plasmids with average sizes of 5.4, 6.8, 7.6, 127, 235 and 359 kb. These replicons showed nucleotide identity greater than 99% with the extrachromosomal elements pTX14-1 (NC_002091), pTX14-2 (NC_004334), pTX14-3 (X56204), pBTHD789-2 (NC_018509), pBtoxis (NC_010076), and pHD1002-1 (NZ_CP009349), respectively (Table 3). The coverage for the plasmid assemblies was between 4000 and 14000 times for the 5.4, 6.8, and 7.6 kb plasmids and between 20 and 109 times for the 127, 235 and 359 kb plasmids.

Table 3. The nucleotide identity between the plasmids of T0124, T0131, T0137 e T013strains and plasmids references.

Plasmids references	pTX14-1	pTX14-2	pTX14-3	pBTHD789-2	pBtoxis	pHD1002-1
Plasmids: Nucleotide identity (%)	pT0124-1: 99.8%	pT0124-2: 99.8%	pT0124-3: 99%	pT0124-4: 99.8%	pT0124-5: 99.9%	pT0124-6: 99.2%
	pT0131-1: 99.8%	pT0131-2: 99.8%	pT0131-3: 99%	pT0131-4: 99.8%	pT0131-5: 99.9%	pT0131-6: 99.2%
	pT0137-1: 99.8%	pT0137-2: 99.8%	pT0137-3: 99.5%	pT0137-4: 99.8%	pT0137-5: 99.9%	pT0137-6: 99.2%
	pT0139-1: 99.8%	pT0139-2: 99.8%	pT0139-3: 99.5%	pT0139-4: 99.8%	pT0139-5: 99.9%	pT0139-6: 99.2%

The 127 kb plasmid is the only one that encodes crystal-forming protein genes that are toxic to Diptera (*cry11Aa*, *cry4Aa*, *cry4Ba*, *cry10Aa*, *cyt1Aa*, *cyt1Ca*, and *cyt2Ba*) (Fig. 3) (Tables 3 and 4). No SNPs were found when the 127 kb plasmids of the different strains were compared.

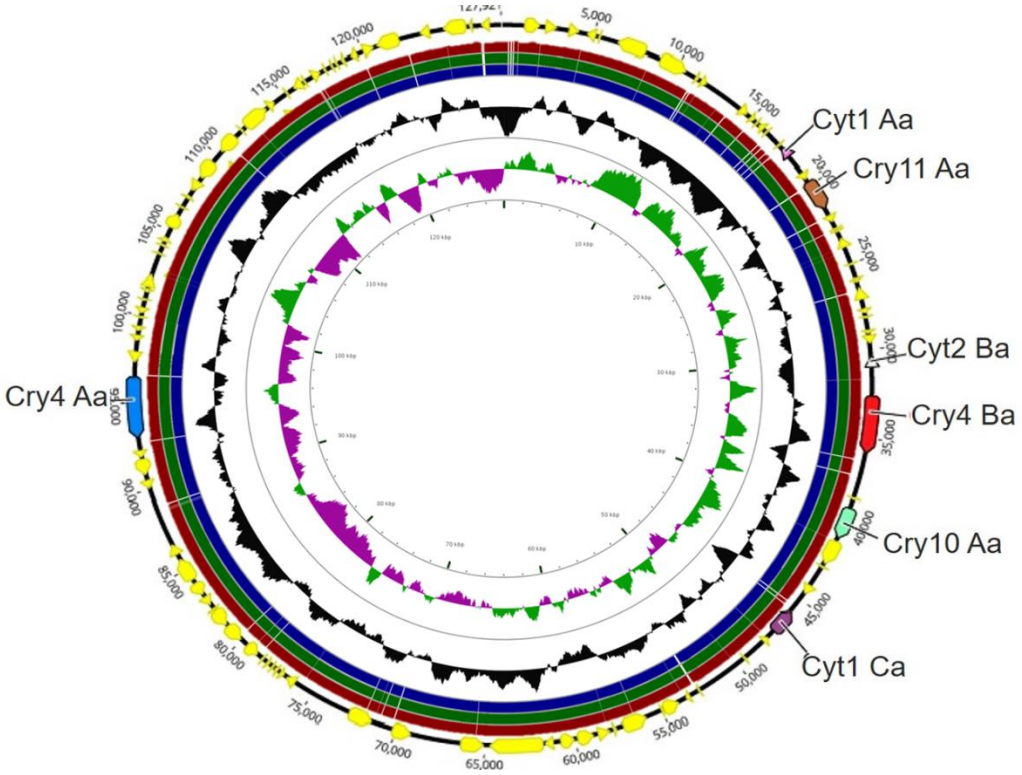

Figure 3. Comparative sequence map of pT0124-4, pT0131-4, pT0137-4 and pT0139-4 plasmids. The circles starting with the outermost ring are as follows: circle 1 (pT0124-4) showing the position of δ-endotoxins; circle 2 (pT0139-4), circle 3 (pT0137-4), and circle 4 (pT0131-4) show regions of sequence similarity representing darker regions detected by BLASTN in the primary sequence (pT0124-4). Circle 5 shows GC content (deviation from average) and circle 6 illustrates the GC skew in green (+) and purple (−). The circle with δ-endotoxins and the map was generated using the Geneious[47] and CGView56 tool.

Table 4. General features of the assembly of complete plasmids of T0124, T0131, T0137 and T0139 strains.

Strains	Plasmids	Average coverage (nº reads)	Standard deviation	Plasmid Size (bp)	GC (%)	CDS	Access number
T0124	pT0124-1	14326	3312	5415	36.	4	CP037884
	pT0124-2	15310	3005	6824	36	3	CP037885
	pT0124-3	5897	982	7697	35.3	9	CP037886
	pT0124-4	109	58.6	127922	32.4	117	CP037887
	pT0124-5	41.5	10	235425	36.6	242	CP037888
	pT0124-6	23.8	14	358206	32.3	338	CP037889
T0131	pT0131-1	14244	2742	5415	36.3	3	CP037453
	pT0131-2	14870	2408	6824	36	3	CP037454
	pT0131-3	5429	794.6	7697	35.3	9	CP037455
	pT0131-4	76.9	41.5	127923	32.4	117	CP037456
	pT0131-5	44.5	8.9	235425	36.6	241	CP037457
	pT0131-6	20.2	11	359437	32.3	336	CP037458
T0137	pT0137-1	8118	2540	5415	36.3	3	CP037459
	pT0137-2	8409	2676	6824	36	3	CP037460
	pT0137-3	3704	832	7697	35.3	9	CP037461
	pT0137-4	124.7	64.4	127923	32.4	117	CP037462
	pT0137-5	55.9	15.4	235425	36.6	241	CP037463
	pT0137-6	33.5	19.5	359440	32.3	336	CP037464
T0139	pT0139-1	4585	48.6	5415	36.3	3	CP037465
	pT0139-2	12014	3 013	6827	36	3	CP037466
	pT0139-3	4423	874.9	7697	35.3	9	CP037467
	pT0139-4	106	55.6	127930	32.3	117	CP037468
	pT0139-5	44	12.1	235425	36.6	241	CP037469
	pT0139-6	28	18.2	359438	32.3	336	CP037470

3.3 Comparative genomic analysis

The genome drafts of the isolates T0124, T0131, T0137, and T0139 were compared to 14 other complete chromosomes of *B. thuringiensis* (Table 5) by phylogenetic analysis and Mauve alignment (Fig. 4).

The Mauve alignment showed collinearity of genes among the isolates from this study and the *Bti* strains AM65-52 and HD-789, forming 32 locally collinear blocks (LCB) (Fig. 4A). The SNP-based phylogeny revealed close relationship between the isolates T0124, T0131, T0137, T0139 and the *Bti* strains AM65-52 and HD-789 (Fig. 4B). Although a total of 2190 SNPs positions were found in all analyzed chromosomes, no SNPs were found in the chromosomes of the isolates used in this study compared to *Bti* strains AM65-52 and HD-789 (Fig. 4C).

Furthermore, the plasmids with 127 kb found in the four isolates (T0124, T0131, T0137, T0139) dif-fered only by minor nucleotide changes (1 to 7 mutations) from the pBtoxis plasmid (NC_010076), and none of the nucleotide changes was related to the δ-endotoxins (Table 6).

Table 5. General features of chromosomes of *Bt* strains used in Mauve alignment and SNP-phylogenetic analysis.

Strain	Status of assembly	Leght (pb)	GC (%)	CDS	Description	Access number	Reference
HS18-1	Complete	5 292 526	35.43	5234	Toxicity to Lepidoptera and Diptera	CP012099.1	Li et al,[56]
MYBT18246	Complete	6 752 490	35.4	6413	Toxicity to nematode	CP015350.1	Unpublished
YC-10	Complete	5 675 007	34.9	6028	Toxicity to nematode	CP011349.1	Cheng et al,[57]
YWC2-8	Complete	5 674 369	35.29	5692	Toxicity to Lepidoptera and Diptera	CP013055.1	Zhu et al,[58]
Bc601	Complete	5 627 121	35.30	5485	Used in fermentation for the production of vitamin C	CP015150.1	Jia et al,[59]
KNU-07	Complete	5 344 151	35.30	5111	Used in agriculture	CP016588.1	Unpublished
Bt185	Complete	5 243 635	35.30	4981	Toxicity to Lepidoptera	CP014282.1	Li et al,[60]
HD1011	Complete	5 232 696	35.5	5245	Medical relevance	CP009335.1	Johnson et al,[61]
HD682	Complete	5 213 295	35. 5	5201	Medical relevance	CP009720.1	Johnson et al,[61]
97-27	Complete	5 235 838	35.4	5216	Medical relevance	CP010088.1	Johnson et al,[61]
HD571	Complete	5 256 240	35.4	5219	Medical relevance	CP009600.1	Johnson et al,[61]
CTC	Complete	5 327 397	35.4	5268	High producer of S-layer protein	CP013274.1	Dong et al,[62]
HD-789	Complete	5 495 278	35.3	5551	Commercial insecticide isolate	CP003763.1	Dogget et al,[21]
AM65-52	Complete	5 499 731	35.0	5463	Toxicity to Diptera	CP013275.1	Bolotin, et al,[24]
T0124	Draft	5 415 530	35.2	5477	Toxicity to Diptera	CP037890	This study
T0131	Draft	5 414 369	35.3	5473	Toxicity to Diptera	CP035735	This study
T0137	Draft	5 414 369	35.3	5472	Toxicity to Diptera	CP035736	This study
T0139	Draft	5 414 367	35.3	5472	Toxicity to Diptera	CP035737	This study

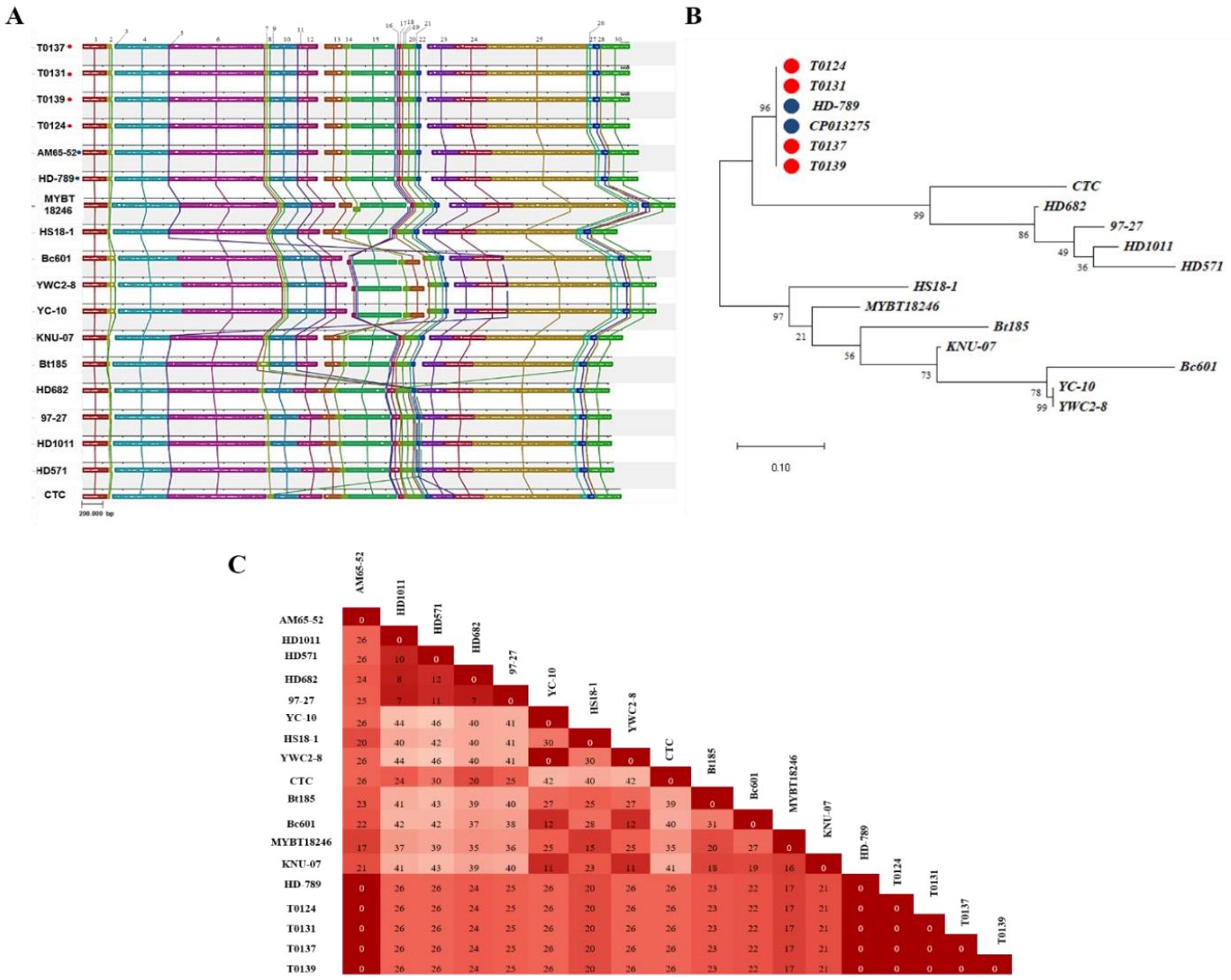

Figure 4. Comparative chromosome sequences of the isolates T0124, T0131, T0137, and T0139 with 14 genomes of other strains of *B. thuringiensis*. (A) Reciprocal LCBs in different sequences are indicated by the same colors and are connected by lines. (B) Phylogenetic tree based on the concatenated SNPs. The SNPs were called by CSI phylogeny 1.452 using HD-789 strain as reference. The branch structure was confirmed by a bootstrap consensus tree inferred from 1,000 replicates in MEGA 10[53]. The scale bar indicates the evolutionary distance between the sequences determined by 0.10 substitutions per nucleotide at the variable positions. Red dots indicate the genomes of strains from the present study and blue dots indicate other genomes of *Bti* from the GenBank database. (C) The matrix shows the 2190 SNPs after pairwise comparison between isolates.

Table 6. The SNPs of the pT0124-4, pT0131-4, pT0137-4, and pT0139-4 using pBtoxis (NC_010076) as reference.

Plasmids	Name	Position	Nucleotide Change	Amino Acid Change	Codon Change	Coverage	Polymorphism Type	Protein Effect	Variant Frequency	Variant P-Value (approximate)
pT0124-4	hypothetical protein CDS	99670	C -> T	S -> F	TCC -> TTC	408	SNP (transition)	Substitution	75.2%	0.0
	hypothetical protein CDS	99354	G -> A	H -> Y	CAT -> TAT	290	SNP (transition)	Substitution	76.9%	0.0
	hypothetical protein CDS	99343	A -> C		GGT -> GGG	304	SNP (transversion)	None	77.6%	7.6E-194
	hypothetical protein CDS	99255	AA -> CC	F -> G	TTT -> GGT	367	Substitution	Substitution	75.2%	0.0
	hypothetical protein CDS	99240	GA -> TG	S -> Q	TCG -> CAG	381	Substitution	Substitution	75.9%	0.0
	hypothetical protein CDS	99234	T -> G		AGG -> CGG	412	SNP (transversion)	None	76.7%	0.0
	hypothetical protein CDS	58097	G -> T		ACC -> ACA	241	SNP (transversion)	None	100.0%	7.9E-25
pT0131-4	hypothetical protein CDS	99670	C -> T	S -> F	TCC -> TTC	238	SNP (transition)	Substitution	79.8%	0.0
	hypothetical protein CDS	99282	T -> G	I -> L	ATA -> CTA	289	SNP (transversion)	Substitution	75.8%	0.0
	hypothetical protein CDS	99255	AA -> CC	F -> G	TTT -> GGT	274	Substitution	Substitution	78.1%	0.0
	hypothetical protein CDS	99240	GA -> TG	S -> Q	TCG -> CAG	282	Substitution	Substitution	78.4%	0.0
	hypothetical protein CDS	99234	T -> G		AGG -> CGG	312	SNP (transversion)	None	80.8%	0.0
	hypothetical protein CDS	99207	T -> C	I -> V	ATT -> GTT	299	SNP (transition)	Substitution	77.6%	0.0
	hypothetical protein CDS	58097	G -> T		ACC -> ACA	183	SNP (transversion)	None	100.0%	5.0E-19
pT0137-4	hypothetical protein CDS	58097	G -> T		ACC -> ACA	323	SNP (transversion)	None	99.7%	4.8E-63
pT0139-4	hypothetical protein CDS	99670	C -> T	S -> F	TCC -> TTC	302	SNP (transition)	Substitution	75.5%	0.0
	hypothetical protein CDS	58097	G -> T		ACC -> ACA	265	SNP (transversion)	None	100.0%	3.2E-27

A functional gene ontology analysis was performed among the four strains (T0124, T0131, T0137, T0139) and two strains of *B. thuringiensis* (HS18-1, and YWC2-8), which presented toxic bioactivity to dipteran insect and not associated to the serotype H14, followed by a summary from shared OrthoVenn clusters. The comparison of the inferred proteins among the strains of this study and the two other strains revealed 4 829 proteins shared by the strains and a total of 231 orthologous clusters shared by HS18-1 and YWC2-8 (Fig. 5A). The HS18-1 and YWC2-8 strains presented specific genes with 6 and 64 single clusters, respectively (Fig. 5A). The analysis of all Gene Ontology (GO) terms assigned to 4 829 orthologous clusters shared by the species showed 1 180 for metabolic processes, 1 001 for ion binding, and 1317 for cell parts (GO-inferred terms) (Fig. 5B–D).

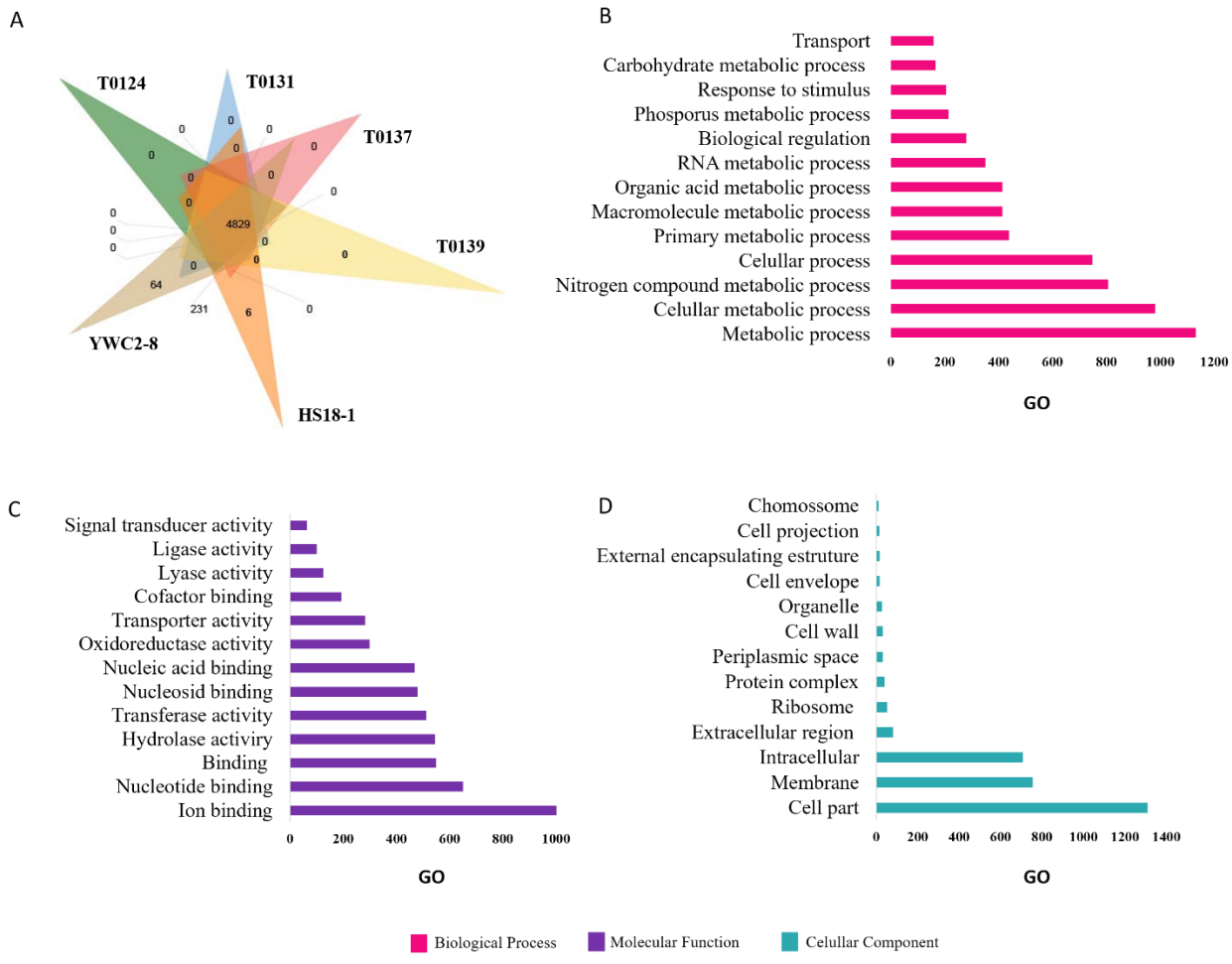

Figure 5. The Venn diagram of the strains from this study and other strains with toxicity to Diptera and summary of the functional gene ontology from shared OrthoVenn clusters. The Venn diagrams of T0124, T0131, T0137, T0139, HS18-1, and YWC2-8 (A). Summary of the functional gene ontology categories using GO slim57 for orthologous clusters in the Venn diagram overlapping regions are represented in the biological process (B), molecular function (C), and cellular component (D) categories.

The important genes of the sporulation process previously described as variable and absent in some *Bacillus* species[25] were analyzed. ORF sequences coding for the germination gene (GerB), small acid soluble proteins genes (SspP and SspH), sensor kinase (SerK) genes, coat gene of the spore (CoatB), and sigma factor genes (SigB, SigE, SigF, and SigH) were compared to the same genes present in other species of the *Bacilli* group (Fig. 6). Higher sequence identity was observed for *B. thuringiensis* HD-789 and *B. thuringiensis* serovar israelensis AM65-52. The GerB and SspH genes showed the highest sequence variability when compared with the sequences acquired in this study.

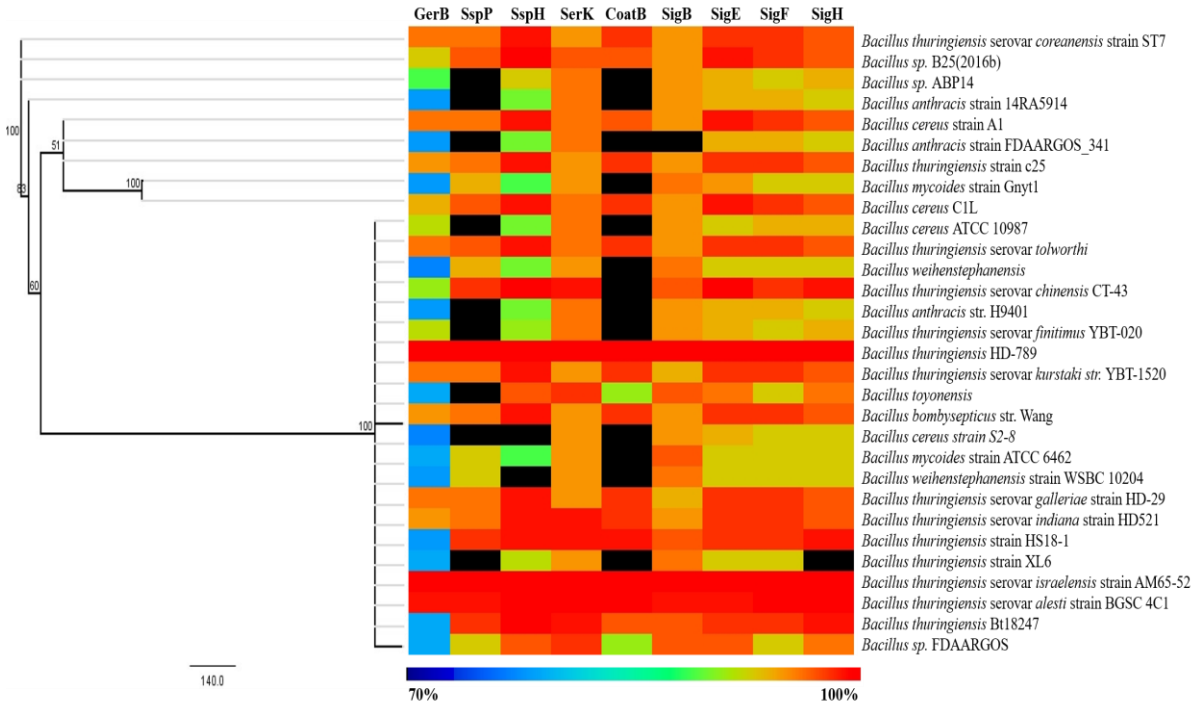

Figure 6. Heatmap comparison of the distribution of sporulation (GerB, SspP, SspH, SerK, CoatB) and sigma factor (SigB, SigE, SigF, and SigH) gene sequences among some species from the *Bacilli* group. Each column and line represent a gene and a *Bacillus* lineage, respectively, and percentage sequence identities between these species and the strains from this study were shown as colors ranging from 70% (dark blue) to 100% (red) as shown in the bottom. Undetected gene sequences are shown in black. The comparative analysis was performed using blastx and the Heatmap was generated using version 4.9.0 of the MeV tool[49].

4 Discussion

Here, the four new *Bti* strains T0124, T0131, T0137, and T0139, collected from the soil of the Tocantins state in Brazil, showed toxic activity to larvae of *A. aegypti* and *C. quinquefasciatus*. In addition to the fact that the mosquito strains were collected from locations not targeted by insecticide applications and hence presenting low risk of insecticide resistance build up.

The *Bti* mode of action is distinct from neurotoxic or growth-regulating compounds used for mosquito control. These facts make *Bti* an effective alternative for controlling mosquito populations displaying or not resistance to these insecticides[26].The *Bti* strains analyzed in this study presented different lethal concentrations among them and when compared to a reference commercial strain of *Bti* (H14). However, the δ-endotoxin gene content and toxin protein profiles assessed by SDS-PAGE were very similar. The T0131 strain presented the highest toxicity for both insect vectors and, therefore, it is probably the most promising strain for biological control among the four isolated strains.

We performed whole genome sequencing of T0124, T0131, T0137, and T0139. However, differently from the toxicity results, the genomic analysis of these isolates indicated highly similar sequences. Previous genome comparison among strains of *Bt* revealed that 80% of the genes of this species are conserved, and the variability among *Bt* strains can be attributed to the acquisition of essential or non-essential genes from other microorganisms residing in the same microbial community[27]. In addition, *Bt* has an open pan-genome which is a characteristic of species that colonize different environments and have different genetic material exchange pathways[28]. *Bt* species comprise different subspecies and comparative analysis of the same subspecies may reveal genomically identical or highly related strains, even from different geographic regions. Such findings could be explained by the emergence of clonal lineages of pathogens that successfully colonized the biosphere, undergoing limited genetic exchange, thus representing homogeneous subspecies[23]. Similarly, studies have shown that *Bti* also present genomically similar strains, indicating the close relationship among them and suggesting a high degree of genomic conservation[29,30] thus corroborating the results obtained in this study.

The assembled chromosomes of T0124, T0131, T0137, and T0139 did not show the presence of the two prophages present in the *Bti* reference HD-78921. Bolotin et al.[24] also identified the sequences the two prophages in the *Bti* AM65-52 strain. Phage sequences, as

well as plasmids, are said to be mobile genetic elements which also contribute to genetic diversity among species and are considered important tools in the divergence of strains and closely related bacterial species[31].

Despite the presence of rearrangements and sequence inversions also that have been linked to the variability of genetically related species[32], the strains from this study, HD-789 and AM65-52 showed collinear chromosomes. Although the polymorphism analysis indicated the presence of various SNPs in the *Bti* isolates, none of the mutations reported was related to the insecticidal activity[33]. The SNP-based phylogeny revealed close relation among the four isolates and other *Bti* isolates (HD-789 and AM65-52), in agreement with previous study[33], reinforcing the close genetic relationship among these bacteria.

With regard to plasmids sequences, the high number of copies of extrachromosomal elements per chromo-some can explain the high coverage of the plasmids with sizes of 5.4, 6.8, and 7.6 kb in this study[24,34]. Although some studies reported the ability of *Bti* strains to harbor up to nine plasmids, the assembly generated here revealed the presence of only six plasmids in the genomes of the isolates, which have also been reported elsewhere[21,24]. The 235 kb plasmids are presented in all sequenced genomes of these *Bti* strains. The 359 kb plasmid was described previously[24] and is also found in the genomes of the T0124, T0131, T0137, and T0139 isolates.

Since the plasmids with 235 kb and 359 kb are not reported to show any crystal-forming proteins linked with toxic activity, they have not been described in this study. In fact, the plasmid with 235 kb has been described as a conjugative plasmid and the plasmid with 359 kb encodes various metabolite transporters[24].

Since the 1980s, the direct relationship between plasmids and the pathogenicity of *Bt* was established, as they are responsible for carrying genes that express toxins active against target insects[35]. Plasmids of 127 kb were found in all isolates of *Bti* containing *cry* and *cyt* genes involved in insect toxicity. This type of plasmid, termed pBtoxis is widely studied and described as the only plasmid capable of encoding the crystal-forming toxins of this bacterium[36]. In addition, pBtoxis also presents sequences with functions predicted to increase crystal formation and subsequent cell viability, acting as chaperones[32,36]. The additional *cry4Ba* coding gene in plasmid pBtoxis-like has been reported to account for the increased effectiveness in mosquito larvae killing of *Bti* strain[33]. In the present study, SNPs analysis of the plasmids carrying the genes encoding the mosquitocidal endotoxins did not reveal any mutation in pT0131-4 what could explain the high toxicity of T0131 strain.

Different strains of *Bt* within the same serotype may share a highly related plasmid pattern; this relationship is most evident in different strains of *Bti* H14 serotype, which, although isolated from different geographic origins, have the same basic plasmid pattern, sometimes even identical[37]. Therefore, our results, that show a high degree of genomic conservation among the strains T0124, T0131, T0137, and T0139, are consistent with previous studies.

The functional gene ontology analysis from shared clusters showed a unique set of proteins identified only in the genome of the YWC2-8 isolate associated with magnesium transport and in the HS18-1 isolate associated with vitamin B6 catabolic processes and pyridoxal 4-dehydrogenase activity (Fig. 5A). The summary of the functional gene ontology showed diversity for metabolic process category (Fig. 5B). The metabolic processes play important roles in the insecticidal activity of *Bt* because metabolic pathways are regulated to provide amino acid, carbon, and energy substances for sporulation and massive synthesis of crystal toxins[38,39].

Heatmap analysis shows that sporulation genes and sigma factors are conserved among *Bti* strains, while the SspH and GerB gene showed highest variability. Although spore formation is central to the definition of *Bacilli*, these genes have been described as variable and absent in some species as a result of niche-specific constraints that may lead to variability in the detection of stress conditions, spore resistance, and germination[25].

The comparative analysis of four new genomes of *Bti* carried out in the present study revealed their very high identity of nucleotide sequence. Furthermore, the results presented here are important for evolutionary studies of this species and potentially may contribute to the improvement of existing strategies or the development of new approaches in biological control that use these bacteria. Further investigations aiming to evaluate potential differences at transcriptomic/proteomic levels during specific phases (e.g., middle vegetative, early sporulation and late sporulation) of the four *Bti* strains will contribute to clarify the higher larvicidal activity described here for the T0131 strain.

5 References

1. Ramírez-Lepe, M. & Ramírez-Suero, M. Biological Control of Mosquito Larvae by *Bacillus thuringiensis* subsp. *israelensis*. (Ed. Perveen, F.), Insecticides-P. Eng., https://doi.org/10.5772/29139 (2012).

2. Liu, G. et al. Complete genome sequence of *Bacillus thuringiensis* subsp. *kurstaki* strain HD73. Genome announc.1, e00080–13, https://doi.org/10.1128/genomea.00080-13 (2013).

3. Barbosa, L. C. B. et al. Draft genome sequence of *Bacillus thuringiensis* 147, a Brazilian strain with high insecticidal activity. Mem. Inst. Oswaldo Cruz.110, 822–823 (2015).

4. He, J. et al. Complete genome sequence of *Bacillus thuringiensis* subsp. *chinensis* strain CT-43. J. Bacteriol.193, 3407–3408 (2011).

5. Palma, L., Muñoz, D., Berry, C., Murillo, J. & Caballero, P. Draft genome sequences of two *Bacillus thuringiensis* strains and characterization of a putative 41.9-kDa insecticidal toxin. Toxins.6, 1490–1504, https://doi.org/10.3390/toxins6051490 (2014).

6. Guan, P. et al. Complete genome sequence of *Bacillus thuringiensis* serovar *sichuansis* strain MC28. J. Bacteriol.194, 6975–6975 (2012).

7. Rusconi, B., Chen, Y., Koenig, S. S., El-Helow, E. R. & Eppinger, M. Sequence of *Bacillus thuringiensis* strain Btm27, an Egyptian isolate highly toxic to cotton leafworm. Genome announc.3, e00446–15, https://doi.org/10.1128/genomeA.00446-15 (2015).

8. Zorzetti, J. et al. Genome sequence of the mosquitocidal *Bacillus thuringiensis* strain BR58, a biopesticide product effective against the coffee berry borer (*Hypothenemus hampei*). Genome announc.3, e01232–15, https://doi.org/10.1128/genomea.01232-15 (2015).

9. Iatsenko, I., Corton, C., Pickard, D. J., Dougan, G. & Sommer, R. J. Draft genome sequence of highly nematicidal *Bacillus thuringiensis* DB27. Genome announc2, e00101–14, https://doi.org/10.1128/genomea.00101-14 (2014).

10. Wang, P. et al. Complete genome sequence of *Bacillus thuringiensis* YBT-1518, a typical strain with high toxicity to nematodes. J. Biotechnol.171, 1–2, https://doi.org/10.1016/j.jbiotec.2013.11.023 (2014).

11. Liu, X. et al. Draft genome sequence of *Bacillus thuringiensis* NBIN-866 with high nematocidal activity. Genome announc.2, e00429–14, https://doi.org/10.3390/toxins6082229 (2014).

12. Wang, A., Pattemore, J., Ash, G., Williams, A. & Hane, J. Draft genome sequence of *Bacillus thuringiensis* strain DAR 81934, which exhibits molluscicidal activity. Genome announc.1, e00175–12, https://doi.org/10.1128/genomea.00175-12 (2013).

13. Jeong, H., Jo, S. H., Hong, C. E. & Park, J. M. Genome sequence of the endophytic bacterium *Bacillus thuringiensis* strain KB1, a potential biocontrol agent against phytopathogens. Genome announc.4, e00279–16, https://doi.org/10.1128/genomeA.00279-16(2016).

14. Loeza-Lara, P. D. et al. The plasmid pBMBt1 from *Bacillus thuringiensis* subsp. *darmstadiensis* (INTA Mo14-4) replicates by the rolling-circle mechanism and encodes a novel insecticidal crystal protein-like gene. Plasmid.54, 229–240, https://doi.org/10.1016/j.plasmid.2005.04.003 (2005).

15. Huang, J., Han, D., Yu, Z. & Sun, M. A novel cryptic plasmid pBMB175 from *Bacillus thuringiensis* subsp. *tenebrionis* YBT-1765. Arch. Microbiol.188, 47–53 (2007).

16. Sheppard, A. E., Poehlein, A., Rosenstiel, P., Liesegang, H. & Schulenburg, H. Complete genome sequence of *Bacillus thuringiensis* strain 407 Cry. Genome announc.1, e00158–12, https://doi.org/10.1128/genomea.00158-12 (2013).

17. Zorzetti, J. et al. Genome sequence of the mosquitocidal *Bacillus thuringiensis* strain BR58, a biopesticide product effective against the coffee berry borer (*Hypothenemus hampei*). Genome announc3, e01232–15, https://doi.org/10.1128/genomea.01232-15 (2015).

18. Berry, C. et al. Complete sequence and organization of pBtoxis, the toxin-coding plasmid of *Bacillus thuringiensis* subsp. *israelensis*. Appl. Environ. Microbiol.68, 5082–5095 (2002).

19. Zghal, R. Z., Ghedira, K., Elleuch, J., Kharrat, M. & Tounsi, S. Genome sequence analysis of a novel *Bacillus thuringiensis* strain BLB406 active against *Aedes aegypti* larvae, a novel potential bioinsecticide. Int. J. Biol. Macromol.116, 1153–1162 (2018).

20. Anderson, I. et al. Comparative genome analysis of *Bacillus cereus* group genomes with *Bacillus suBtilis*. FEMS Microbiol. Lett.250, 175–184 (2005).

21. Doggett, N. A. et al. Complete genome sequence of *Bacillus thuringiensis* serovar *israelensis* strain HD-789. Genome announc.1, e01023–13, https://doi.org/10.1128/genomea.01023-13 (2013).

22. Jeong, H., Seung-Hwan, P. & Soo-Keun, C. Genome sequence of the acrystalliferous *Bacillus thuringiensis* serovar israelensis strain 4Q7, widely used as a recombination host. Genome announc2, e00231–14, https://doi.org/10.1128/genomea.00231-14 (2014).

23. Ankarloo, J. et al. Genome stability of *Bacillus thuringiensis* subsp. *israelensis* isolates. Curr. Microbiol.40, 51–56 (2000).

24. Bolotin, A. et al. Comparative genomics of extrachromosomal elements in *Bacillus thuringiensis* subsp. *israelensis*. Microbiol. Res.168, 331–344, https://doi.org/10.1016/j.resmic.2016.10.008 (2016).

25. Alcaraz, L. D. et al. Understanding the evolutionary relationships and major traits of *Bacillus* through comparative genomics. BMC genomics.11, 332, https://doi.org/10.1186/1471-2164-11-332 (2010).

26. Araújo, A. P. et al. The susceptibility of *Aedes aegypti* populations displaying temephos resistance to *Bacillus thuringiensis israelensis*: a basis for management. Parasit vectors.6, 297, https://doi.org/10.1186/1756-3305-6-297 (2013).

27. Fang, Y. et al. A pangenomic study of *Bacillus thuringiensis*. J. Genet. Genomics.38, 567–57 (2011).

28. Medini, D., Donati, C., Tettelin, H., Masignani, V. & Rappuoli, R. The microbial pan-genome. Curr. Opin. Genetics Dev.15, 589–594 (2005).

29. Rivera, A. M. G. & Priest, F. G. Molecular typing of *Bacillus thuringiensis* serovars by RAPD-PCR. Syst. Appl. Microbiol.26, 254–261 (2003).

30. Sorokin, A. et al. Multiple-locus sequence typing analysis of *Bacillus cereus* and *Bacillus thuringiensis* reveals separate clustering and a distinct population structure of psychrotrophic strains. Appl. Environ. Microbiol.72, 1569–1578 (2006).

31. Frost, L. S., Leplae, R., Summers, A. O. & Toussaint, A. Mobile genetic elements: the agents of open-source evolution. Nat. Rev. Microbiol.3, 722–732, https://doi.org/10.1038/nrmicro1235 (2005).

32. Rasko, D. A., Altherr, M. R., Han, C. S. & Ravel, J. Genomics of the *Bacillus cereus* group of organisms. FEMS Microbiol. Lett.29, 303–329 (2005).

33. Fayad, N. et al. Characterization and Whole Genome Sequencing of AR23, a Highly Toxic *Bacillus thuringiensis* Strain Isolated from Lebanese Soil. Curr. Microbiol.76, 1503–1511 (2019).

34. Galvão, P., Gitahy, P. D. M., Vidal, M., Araujo, J. L. S. & Baldani, J. Plasmídios em *Bacillus thuringiensis*: tamanho, mecanismo de replicação e papel na atividade entomopatogênica. Embrapa Agrobiologia- (Infoteca-e).262, 1–32 (2009).

35. González, J. M., Dulmage, H. T. & Carlton, B. C. Correlation between specific plasmids and δ-endotoxin production in *Bacillus thuringiensis*. Plasmid.5, 351–365, https://doi.org/10.1016/0147-619X(81)90010-X (1981).

36. Stein, C., Jones, G. W., Chalmers, T. & Berry, C. Transcriptional analysis of the toxin-coding plasmid pBtoxis from *Bacillus thuringiensis* subsp. *israelensis*. Appl. Environ. Microbiol.72, 1771–1776 (2006).

37. Reyes-Ramírez, A. & Ibarra, J. E. Plasmid patterns of *Bacillus thuringiensis* type strains. Appl. Environ. Microbiol.74, 125–129 (2008).

38. Deng, C., Peng, Q., Song, F. & Lereclus, D. Regulation of cry gene expression in *Bacillus thuringiensis*. Toxins.6, 2194–2209, https://doi.org/10.3390/toxins6072194 (2014).

39. Wang, J. et al. The metabolic regulation of sporulation and parasporal crystal formation in *Bacillus thuringiensis* revealed by transcriptomics and proteomics. Mol. Cell Proteomics.12, 1363–1376 (2013).

40. Monnerat, R. G., Silva, S. F. & Silva-werneck, J. O. Catálogo do banco de germoplasma de bactérias entomopatogênica do gênero *Bacillus*. Documentos, Brasília: Embrapa-Cenargen.60, 65, https://doi.org/10.1016/j.biocontrol.2006.11.008 (2001).

41. Lecadet, M. M., Chaufaux, J., Ribier, J. & Lereclus, D. Construction of novel *Bacillus thuringiensis* isolates with different insecticidal activities by transduction and transformation. Appl. Environ. Microbiol.58, 840–849 (1992).

42. Mounsef, J. R. et al. A simple method for the separation of *Bacillus thuringiensis* spores and crystals. J. Microbiol. Methods.107, 147–149 (2014).

43. Sambrook, J. & Russel, D. W. Molecular cloning: A laboratory manual, (Ed 3 Lab) Cold Springs Harbour Press. 2344, https://doi.org/10.1016/0307-4412(83)90068-7 (2001).

44. Aguiar, R. W. S., Martins, E. S., Ribeiro, B. M. & Monnerat, R. G. Cry10Aa protein is highly toxic to *Anthonomus grandis* Bheman (Coleoptera: Curculionidae), an important insect pest in Brazilian cotton crop fields. Bt Research.3, 20–28, https://doi.org/10.5376/bt.2012.03.0004 (2012).

45. McLaughlin, R. E. et al. US standard bioassay for the potency assessment of *Bacillus thuringiensis* serotype H-14 against mosquito larvae. Bulletin of the ESA. 30, 26–29, https://doi.org/10.1093/besa/30.1.26 (1984).

46. Andrews S. FastQC: a quality-control tool for high-throughput sequence (2015).

47. Kearse, M. et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics.28, 1647–1649 (2012).

48. Alikhan, N. F., Petty, N. K., Zakour, N. L. B. & Beatson, S. A. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. BMC genomics12, 402, https://doi.org/10.1186/1471-2164-12-402 (2011).

49. Saed, A. I. et al. TM4: a free, open-source system for microarray data management and analysis. Biotechniques.34, 374–378 (2003).

50. Wang, Y., Coleman-Derr, D., Chen, G. & Gu, Y. Q. OrthoVenn: a web server for genome wide comparison and annotation of orthologous clusters across multiple species. Nucleic Acids Res.43, W78–W84 (2015).

51. Darling, A. C., Mau, B., Blattner, F. R. & Perna, N. T. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res.14, 1394–1403 (2004).

52. Kaas, R. S., Leekitcharoenphon, P., Aarestrup, F. M. & Lund, O. Solving the problem of comparing whole bacterial genomes across different sequencing platforms. Plos One.9, e104984, https://doi.org/10.1371/journal.pone.0104984 (2014).

53. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Mol Biol Evol.35, 1547–1549 (2018).

54. SAS Institute. SAS/STAT User's Guide., Cary, NC, USA, (2008). 55. Robertson, J. L., Jones, M. M., Olguin, E., & Alberts, B. Pesticide Bioassays with Arthopods. Boca Raton: CRC Press. 224 p. (2007).

56. Stothard, P. & Wishart, D. S. Circular genome visualization and exploration using CGView. Bioinformatics.21, 537–539 (2004).

57. Gene ontology consortium. The Gene Ontology (GO) database and informatics resource. Nucleic Acids Res.32, D258–D261 (2004).

58. Li, Q. et al. Complete genome sequence of *Bacillus thuringiensis* HS18-1. J. Biotechnol.214, 61–62 (2015). 59. Cheng, F. et al. Complete genome sequence of *Bacillus thuringiensis* YC-10, a novel active strain against plant-parasitic nematodes. J. Biotechnol.210, 17–18 (2015).

60. Zhu, J. et al. The complete genome sequence of *Bacillus thuringiensis* serovar *hailuosis* YWC2-8. J. Biotechnol.219, 38–39 (2016).

61. Jia, N., Ding, M. Z., Gao, F. & Yuan, Y. J. Comparative genomics analysis of the companion mechanisms of *Bacillus thuringiensis* Bc601 and *Bacillus endophyticus* Hbe603 in bacterial consortium. Sci. Rep.6, 28794, https://doi.org/10.1038/srep28794 (2016).

62. Li, Y. et al. Complete genome sequence of *Bacillus thuringiensis* Bt185, a potential soil insect biocontrol agent. J. Integr. Agric.16, 749–751 (2017).

63. Johnson, S. L. et al. Complete genome sequences for 35 biothreat assay-relevant *Bacillus* species. Genome announc.3, e00151–15, https://doi.org/10.1128/genomeA.00151-15 (2015).

64. Dong, Z. et al. Complete genome sequence of *Bacillus thuringiensis* CTC—A typical strain with high production of S-layer proteins. J. Biotechnol.220, 100–101 (2016).

CHAPTER II: Isolation, genomic, and proteomic characterization of a novel Neotropical strain of *Bacillus thuringiensis* with mosquitocidal activities

Submitted to: Processes (Annex B)

Abstract: The combination of genomic and proteomic analysis are useful tools for the study novel *Bacillus thuringiensis* (Bt) strains, as these approaches allow the accurate identification of pesticidal proteins and virulence factors produced. Here, we isolated and evaluated the potential a novel Neotropical *Bt* strain (TOD651) for controlling larvae of *Aedes aegypti* and *Culex quinquefasciatus* mosquitoes. Aiming the full comprehension of the TOD651 larvicidal potential, we further evaluated the whole TOD651 genome and conducted the proteomic analysis of the TOD651 spore-crystal mixtures. Our results showed that *Bt* TOD651 similarly killed both *A. aegypti* (0.011 µg/mL) and *C. quinquefasciatus* (0.023 µg/mL) larvae, exhibiting similar potency to the commercial *Bt* strain. The genome sequence revealed that *Bt* TOD651 harboring *cry11Aa3*, *cry10Aa4*, *cry4Aa4*, *cry4Ba5*, *cyt1Aa5*, *cyt1Ca1*, *cyt2Ba13*, *mpp60Aa3*, and *mpp60Ba3*. The proteomic analysis revealed not expression of Mpp60Aa3, while all the other pesticidal proteins were expressed (Cry4Ba5 was more abundant that Cyt1Aa5). The expression of the Mppe showed the major proportions between proteases. The virulent factor Neutral protease B and Spore coat proteins also were expressed. The expression of relevant pesticidal proteins (e.g., Cry, Cyt, Mpp, and other pathogenic factors), whose actions can occur in a synergic relation, indicate the biocontrol using *Bt* TOD651 may contribute to delay the selection of resistant individuals.

Keywords: *Bacillus thuringiensis*, Genome sequencing, Proteomic, Biorational mosquito control, Pesticidal proteins.

1 Introduction

The control of insect vectors of different diseases is of great importance for public health. Mosquito species such as *Aedes aegypti* (Linnaeus, 1762) (Diptera: Culicidae) and *Culex quinquefasciatus* (Say, 1823) (Diptera: Culicidae) can transmit several diseases that affect human life. For instance, *A. aegypti* can transmit yellow fever virus, dengue virus, Chikungunya virus, and Zika virus; while the *C. quinquefasciatus* is capable of transmitting arboviruses like West Nile Virus (WNV) and the Wuchereria bancrofi nematode, responsible for the lymphatic filariasis disease [1].

The use of chemical insecticides for the control of mosquitoes, which harm the environment, has been slowly substituted around the world by biological control strategies such as *Bacillus thuringiensis* (*Bt*) [2]. Bt is a gram-positive bacterium known for its toxicity and specificity towards insect hosts due to its ability to produce and release crystal pro-teins (Cry and Cyt) during the sporulation stage [3]. *Bacillus thuringiensis* serovar *israelensis* (*Bti*) is one of the subspecies' most effective larvicides for mosquito control, being recom-mended by the World Health Organization (WHO) [4,5]. Bti produces Cry and Cyt crystals (Cry4Aa, Cry4Ba, Cry10Aa, Cry11Aa, Cyt1Aa, and Cyt2Ba) that exhibit toxicity against mosquito species from the genus *Aedes, Anopheles, and Culex*, and for the black fly (Sim-uliidae) [4, 6, 7]. These proteins interact synergistically and can decrease the incidence of resistance in insect populations [4]. Some *Bti* strains can also harbor Mpp60A and Mpp60B proteins, which are present in other subspecies such as *jegathesan* and *malayensis* [9].

Improvements in the control of mosquitoes using *Bt* have involved the constant identification and characterization of novel strains and pesticidal proteins [9-12]. In addition, next-generation sequencing (NGS) has allowed the whole-genome sequencing of novel *Bt* strains and their characterization. Genome information has been important in research and applications of Bt because pesticidal genes are easily detected [13-15]. Furthermore, other genes related to the pathogenicity of *Bt* can be explored through genome sequencing, such as virulence factors and other secondary metabolites [16,17].

It is worth noting that despite the characterization of genes coding pesticidal proteins allowing strain classification, it is the expression of these genes that determines their spectrum of activity [18,19]. In the genome sequence, not all coding regions annotated are expressed, and the genomic approach may not suffice to fully explain toxicity differences between *Bt* strains [20,21]. For instance, many pesticidal proteins are cryptic or with in-significant levels of expression [22]. Thus, proteomic analysis of the pesticidal proteins that make up the

parasporal crystal is essential to understand the toxicity of novel and commercial *Bt* strains [19]. In this context, the combination of genomic and proteomic analysis is a powerful tool for the accurate identification of pesticidal proteins and viru-lence factors of *Bt* strains [23-25], and estimations of the abundance of such proteins can be achieved in purified parasporal crystals and spore-crystal mixtures [25,26].

Several studies have been conducted to investigate the genomics of mosquitocidal *Bt* strains [21,27-29]. Interestingly, the proteomic analysis for pesticidal proteins responsible for mosquitocidal activities in *Bt* strains remains scarce and underexploited.

Therefore, we isolated a novel *Bt* strain (*Bt* TOD651) and evaluated its insecticidal activity against larvae of *A. aegypti* and *C. Quinquefasciatus*. In order to explore and better understand its toxicity, we sequenced the whole genome of TOD651 and performed proteomic analysis of the spore-crystal mixture.

2 Materials and Methods

2.1 Origin and culture of Bt TOD651 strain

The *Bt* TOD651 was isolated from a soil sample, collected in the state of Tocantins, Brazil (11º43'45'' S; 49º04'07" W), according to Monnerat et al. (2007) [30]. This strain was cultured at 28 °C for 12 h in Luria-Bertani (LB) solid medium (10 gL^{-1} tryptone, 5 gL^{-1} yeast extract, 10 gL^{-1} NaCl, and 20 gL^{-1} agar). Posteriorly, a single colony of *Bt* TOD651 was transferred to an LB liquid medium and incubated (28 °C at 200 rpm for 16 h) for sporulation and DNA extraction steps. The *Bti* AM65-52 was isolated from a commercial sample (VectoBac®, Sumitomo) and used as a reference strain.

2.2 Crystal protein purification and SDS–PAGE analysis

An aliquot of LB culture (3 mL) was transferred to CCY medium (30 mL) [13 mM KH2P04, 26 mM K2HP04, 0.002% (w/v) L-glutamine, 0.1% (w/v) casein hydrolysate, 0.1 % (w/v) bacto casitone, 0.04% bacto yeast extract, 0.6% (w/v) glycerol, 0.05 M $ZnCl_2$, 0.5 M $MgCl_2$, 0.01 M $MnCI_2$, 0.2 M $CaCl_2$, 0.05 M $FeCl_3$] and incubated for sporulation (28 °C at 200 rpm for 72h). Then, the spore-crystal mixture was collected, and the crystal proteins were purified according to a previously described method [31]. Purified crystals were suspended in a small volume of phosphate-buffered saline (136 mM NaCl, 1.4 mM KH_2PO_4, 2.6 mM KCl, 8 mM Na_2HPO_4, and 4.2 mL H_2O; pH 7.4), and fractionated by electrophoresis on a 10% SDS-PAGE gel [30].

2.3 Identification of crystal morphology

The morphological characterization of Cry protein crystals was performed by scanning electron microscopy. The spore-crystal mixture of *Bt* TOD651 was collected and diluted in sterile water. Then, 100 µL of the diluted suspension was placed on metallic supports and dried for 24 h at 37 °C, covered with gold for 180 s using an Emitech apparatus (model K550; Quorum Technologies, Lewes, UK), and observed under a Zeiss scanning electron microscope (model DSM 962; Carl Zeiss AG, Oberkochen, Germany) at 10 or 20 Kv.

2.4 Larvae rearing

The larvae of *A. aegypti* and *C. quinquefasciatus* were collected from fields without the application of insecticides, in regions of transition between urban and rural areas in the state of Tocantins, Brazil (11°40'55.7" latitude S, 49°04'3.9" longitude W). The insect colonies were established in the Entomology Laboratory of the Federal University of Tocantins, Gurupi Campus, according to Aguiar et al. [32]. The larvae were reared in plastic containers (40cm x 25cm x 8cm) and fed a sterilized diet (an 80/20 mix of chick chow powder/yeast), and mosquitoes were provided with a 10% sucrose solution and the blood of live Wistar rats (*Rattus norvegicus albinus*). The guidelines (international, national, and institutional) for the care and use of animals were followed.

2.5 Bioassays

Bioassays were conducted on *A. aegypti* and *C. quinquefasciatus* third-instar larvae using spore-crystal mixtures. The concentrations were determined according to Mclaughlin et al. [33]. Seven concentrations of the spore-crystal mixtures (0.05, 0.10, 0.15, 0.20, 0.25, 0.30, and 0.40 µg/ml) were tested, and sterile distilled water was used as a negative control. Bioassays were performed in three replicates with 25 larvae in 100 mL of distilled water. Treated larvae were kept at 26 ± 1 °C, 60.0 ± 5% RH, 12 h light-dark photoperiod for 24h before being examined. The spore-crystal mixture from the AM65-52 strain was used as a reference. Concentration–mortality curves were estimated by probity analysis using the PROBIT procedure in the SAS software [34].

2.6 Whole-genome sequencing, assembly, and annotation

Genomic DNA was extracted using the Wizard® Genomic DNA Purification Kit (Promega, Madison, WI, USA). Posteriorly, DNA concentration, and purity were measured using the NanoDrop™ 8000 apparatus (Thermo Fisher Scientific, Waltham, MA, USA) and

stored at -20 °C until further use. Sequencing was performed on Illumina Mi-Seq technologies (paired-end application, reads with a mean length of 75.9 bp) (Illumina, San Diego, CA, USA), generating a total of 15,425,426 reads with an average insert size of 200 bp and coverage of 426X. Sequence reads quality was assessed using FastQC software version 0.11.9 [35] and reads were trimmed using the Trim and Filter tool (Error proability= 0.05) of Geneious version 10.2.6 [36]. The trimmed reads were used in de novo assembly with the SPAdes version 3.10.0 tool and default parameters [37] and contigs ≥1000 bp were discarded. The CDS of contigs were predicted using RASTtk (Domain: Bacteria; Taxonomy name: *Bacillus thuringiensis*; Genetic code: 11 - Archaea and Bacteria). The chromosome was assembled using contigs and reference HD-789 (NCBI accession n° CP003763) through reference-guided *de novo* assembly [38] using Geneious' map to reference tool to assess the virulence factors of related genes. Contigs unused in the chromosome assembly were filtered and used for predicting pesticidal protein-like genes. Related genes with virulence factors were predicted using the bacterial virulence factor database (VFDB) [39]. Putative pesticidal proteins were determined using Blastx through the Btoxin_Digger tool (scaffolds as a query) [40] and a customized database (CDS predicted as a query). The customized database was created from the *Bt* pesticidal protein list available at the *Bt* nomenclature website (http://www.lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/toxins2.html) through Geneious using the Add/Remove Database tool. CDS with homology to the *Bt* pesticidal proteins were filtered using E-value 0.001 and word size 6 parameters.

2.7 Phylogenetic relationship

A phylogenetic tree was constructed using the *gyrB* gene (DNA gyrase subunit B), extracted from contigs sequences and the *gyrB* genes of the *Bacillus* ssp. strains retrieved from GenBank. The alignment was performed using ClustalW, and the phylogenetic tree was created using MEGA 11 [41] from the neighbor-joining method with 1000 replications.

2.8 LC-MS/MS analysis

The liquid chromatography–tandem mass spectrometry (LC-MS/MS) method was used for protein detection in the spore-crystal mixture of the *Bt* TOD651 strain. The LC-MS/MS analysis was carried out at the Veritas/Life Sciences at the University of São Paulo (USP, Ribeirão Preto, SP, Brazil). Firstly, the spore-crystal sample was washed three times in 1X PBS (Phosphate-Buffered Saline), resuspended in 750 µL of solubilization buffer (8M urea,

0.5% Octyl-glucopyranoside (OG), 0.05M Tris-HCL, pH 8.8), and sonicated by three cycles (60 seconds, 30% amplitude, and shut off for two seconds) while maintained on ice. The quantification of solubilized protein was performed using the Bradford method (Protein Assay Dye Reagent Concentrate, Bio-Rad Laboratories) according to the manufacturer's instructions. In the sample preparation for advanced mass spectrometry, 50 µg of sample was subjected to disulfide bridge reduction (50µg of DTT [Dithiothreitol]; 60 min of incubation at 37 °C), followed by alkylation (250 µg of I.A [iodocetamide]; 60 min at room temperature in the dark). Following this, the sample was diluted five times in Tris hydrochloride (0.05M Tris-HCl, pH 8.8) and incubated using 2 µg of trypsin (Promega, V511A) at 37 °C overnight. The cleanup and desalting of the sample were performed using C18 resin (Supleco). The column was calibrated using 2% acetonitrile containing 0.1% formic acid, and the elution was performed with 50% acetonitrile. The sample was then dried in a speed vac and applied to a mass spectrometer (Thermo Fisher Orbitrap Eclipse) coupled to a nanoflow Nano LC-MS/MS chromatography system (Dionex Ultimate 3000 RLSCnano System, Thermofisher). Peptides were separated for 90 minutes in a nanoEase MZ peptide BEH C18 column (130A, 1.7 µm, 75 µm x 250 mm, Waters) at 300 nL/min with a 4-50% acetonitrile gradient. The data were obtained on MS1 in the range of M/Z 375-1500 (120,000 resolution, AGC target 1E6, maximum time injection of 100 ms). The most abundant ions were submitted to MS/MS (30% collision energy, 1.2 m/z, AGC target 1E5, 15000 resolution).

2.9 Proteomic data analysis

The proteomic data was processed using PatternLabV [42]. Firstly, the customized database was created using translated CDS of the TOD651 genome through the Generate Search DB option, including a contaminant library (MS contaminant sequences, e.g., trypsin, keratins, and albumin). Then, proteomic data was analyzed against the customized database using the following parameters: The modifications selected in the search were carbamidomethyl (C), deamination (NQ), and oxidation (M). Enzyme trypsin (fully specific), two maximum missed cleavages, an initial precursor mass tolerance of 35 ppm, MS and MS/MS tolerance errors of 10 ppm, and acceptable FDR (False Discovery Rate) estimates of 3% at spectral, 2% at peptide, and 1% at protein levels were added as advanced parameters.

The functional annotation of the identified proteins was performed with the UniProtKB/Swiss-Prot database, and the summary graphical of functional classification was

created using GO terms through the WEGO 2.0 tool (Web Gene Ontology Annotation Plot) [43].

2.10 Data availability

The clean reads of the *Bt* TOD651 have been deposited at the Sequence Read Archive (SRA) under the accession number PRJNA907848.

3 Results

3.1 Protein profile, crystals morphology, and mosquitocidal activity

The protein profile of *Bt* TOD651 purified crystals in an SDS-PAGE showed main proteins with molecular weights of approximately 130, 70, and 27 kDa size (Figure 1a). The ultra-structural an b). alysis of the spore-crystal mixture indicated the presence of spherical crystals (Figure 1).

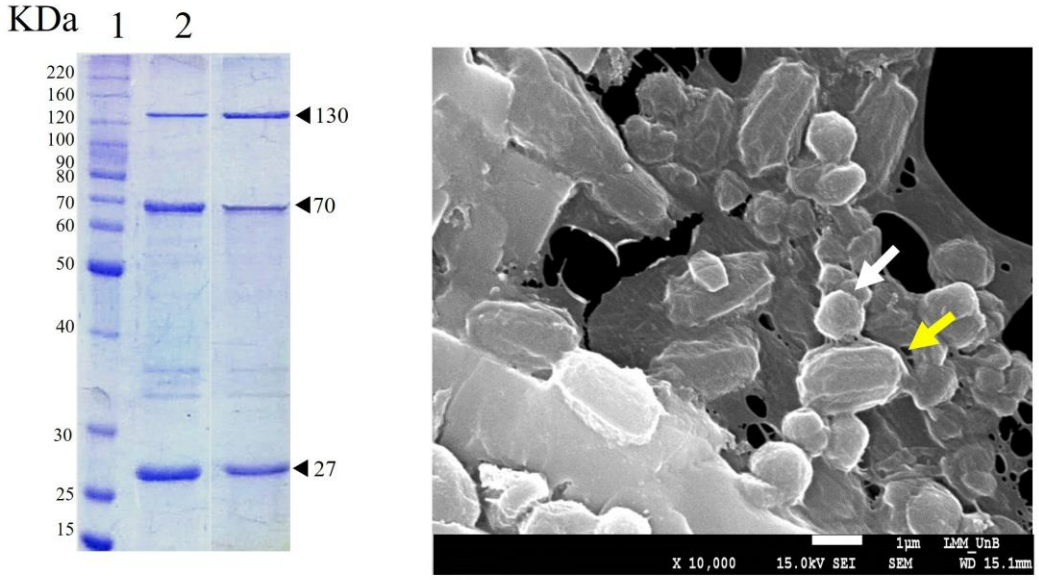

Figure 1. SDS-PAGE analysis and scanning electron microscopic of *Bt* TOD651 purified crystals and spore-crystal mixture. **(a)** Protein profile of *Bt* TOD651 purified crystals: Lane 1 - molecular mass marker; Lane 2 – AM65-52 purified crystals; Lane 3 - TOD651 purified crystals. **(b)** Scanning electron micrograph of spores-crystals mixture *Bt* TOD651. Arrows indicate spore (yellow) and crystal (white).

The TOD651 spore-crystal mixture showed mosquitocidal activity towards *A. aegypti* and *C. quinquefasciatus*. The high toxicity was observed, but the lethal concentration did not differ between the reference strain and TOD651 for both species; however, TOD651 showed significantly higher toxicity to *A. aegypti* when compared to *C. quinquefasciatus*, with 50% lethal concentration (LC_{50}) values of 0.011 and 0.023 µg/mL, respectively (Table 1).

Table 1. Lethal concentrations estimations of *Bt* TOD651 to larvae of *Aedes aegypti* and *Culex quinquefasciatus*.

Strain	*A. aegypti*				*C. quinquefasciatus*				
	LC_{50} (µg/mL)	**CL_{95} (µg/ml)**	**SLOPE**	**χ^2**	**LC_{50} (µg/mL)**	**LC_{95} (µg/mL)**	**SLOPE**	**χ^2**	***P***
TOD651	0.011	0.030	3.726	5.62	0.023	0.055	4.311	6.68	
AM65-52	0.013	0.037	3.725	4.33	0.028	0.069	4.467	6.49	

3.2 General genomic features

The majority of genes were classified into different functional classes using the RASTtk tool's analysis. The amino acids and derivatives metabolism class (579 genes), carbohydrate metabolism (320 genes), cofactors, vitamins, prosthetic groups, pigment metabolism subsystems (215 genes), protein metabolism (155 genes), cell wall and capsule (163), nucleosides and nucleotides (159 genes), and protein metabolism were the most common (Figure 2). One hundred and eight (108) genes were grouped in the subsystem class "virulence, disease, and defense."

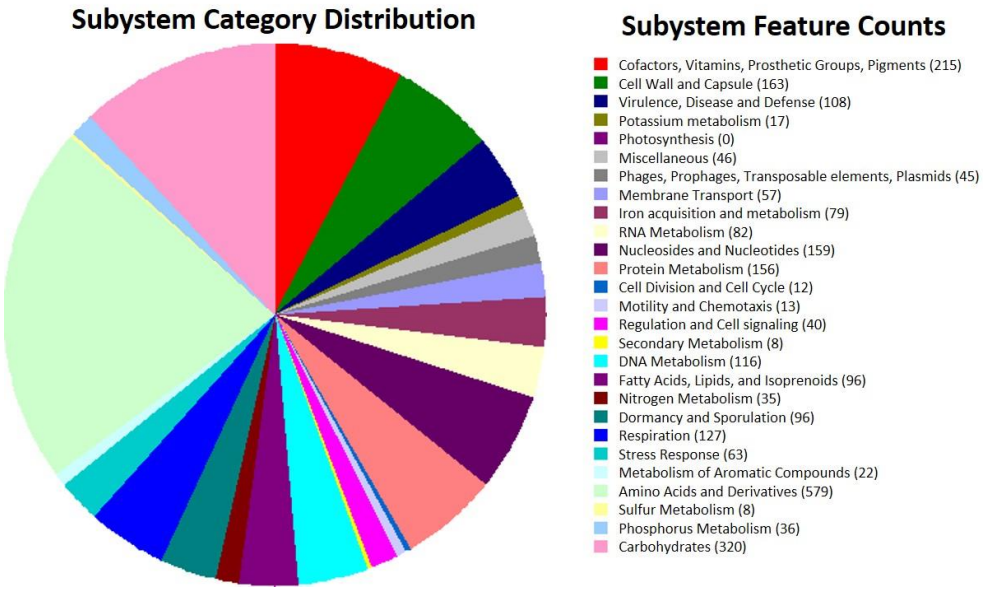

Figure 2: Subsystem category distributions in genome of *Bt* TOD651 based functional classification.

To assess the virulence factors, the chromosome was assembled at the draft level. This sequence consisted of a chromosome with ~ 5.4 Mb bp containing 35.9% GC, 5130 CDS, 1 rRNA, and 71 tRNA genes (Table 2). The sequences not used in the chromosome, and presumably belonging to plasmid sequences, were used for the *cry*/*cyt* gene screening.

Table 2. Draft chromosome features of *Bt* TOD651 strain.

General features	**Value**
Mean coverage (nº reads)	95.9
Chromosome size (bp)	5,409,948
Gapped sites (%)	7.2
GC content (%)	35.9
CDS (n°)	5,130
rRNA(n°)	1
tRNA (n°)	71

3.3 Phylogenetic analysis

Phylogenetic analysis using the *gyrB* gene showed that the *Bt* TOD651 strain forms a group closely related to three *Bti* strains (BGSC 4Q1, BGSC 4Q7rifR, and AM65-52), *Bt* MYBT18246, *Bt* ATCC 10792, and *Bt* serovar *thuringiensis* IS5056 (Figure 3).

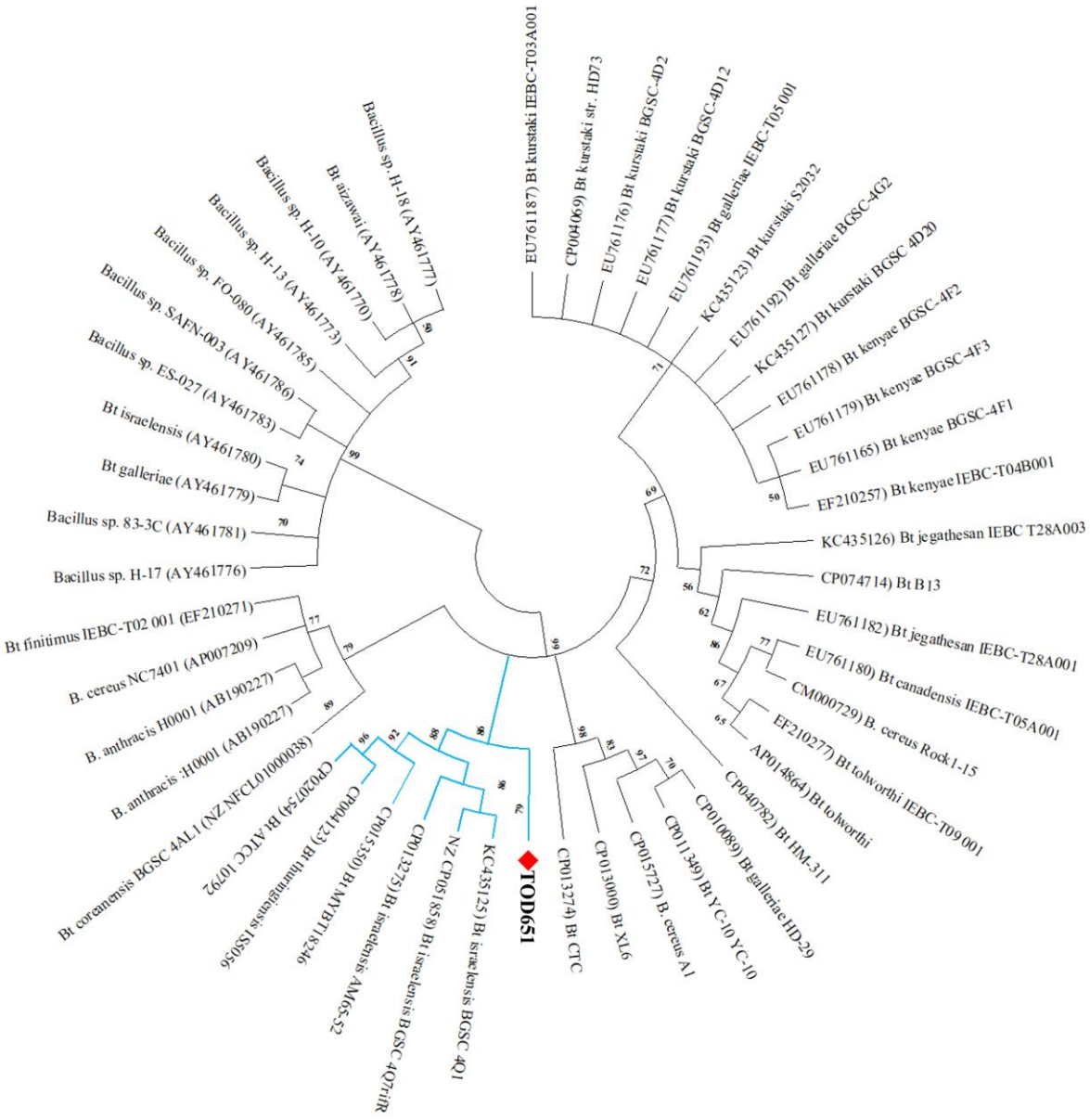

Figure 3. Phylogenetic analysis of *Bt* TOD651 with other *Bacillus* spp. based on *gyrB* sequences. The phylogenetic tree was constructed by using the Neighbor-Joining method with 1000 bootstrap replications. Bootstrap values <50% were disregarded for branches.

3.4 Genes associated with Bt TOD651 pathogenicity

Different virulence-associated genes were detected in the genome sequence (Table S1, Figure 4). Among them, enzymes (*inhA1* and *sph*), immune evasion (*bpsC* and polysaccharide capsule genes), iron acquisition genes (*dhbA-C*, *dhbE*, *dhbF*, *hal*, *ilsA*, and *asbA-F*), regulation genes (*pagR-XO2*, *papR*, *plcR*, and *cheA*), and toxins (*hlyI-III*, *hblA*, and *nheA-C*) were identified (Supplementary Table S1, Figure 4 a).

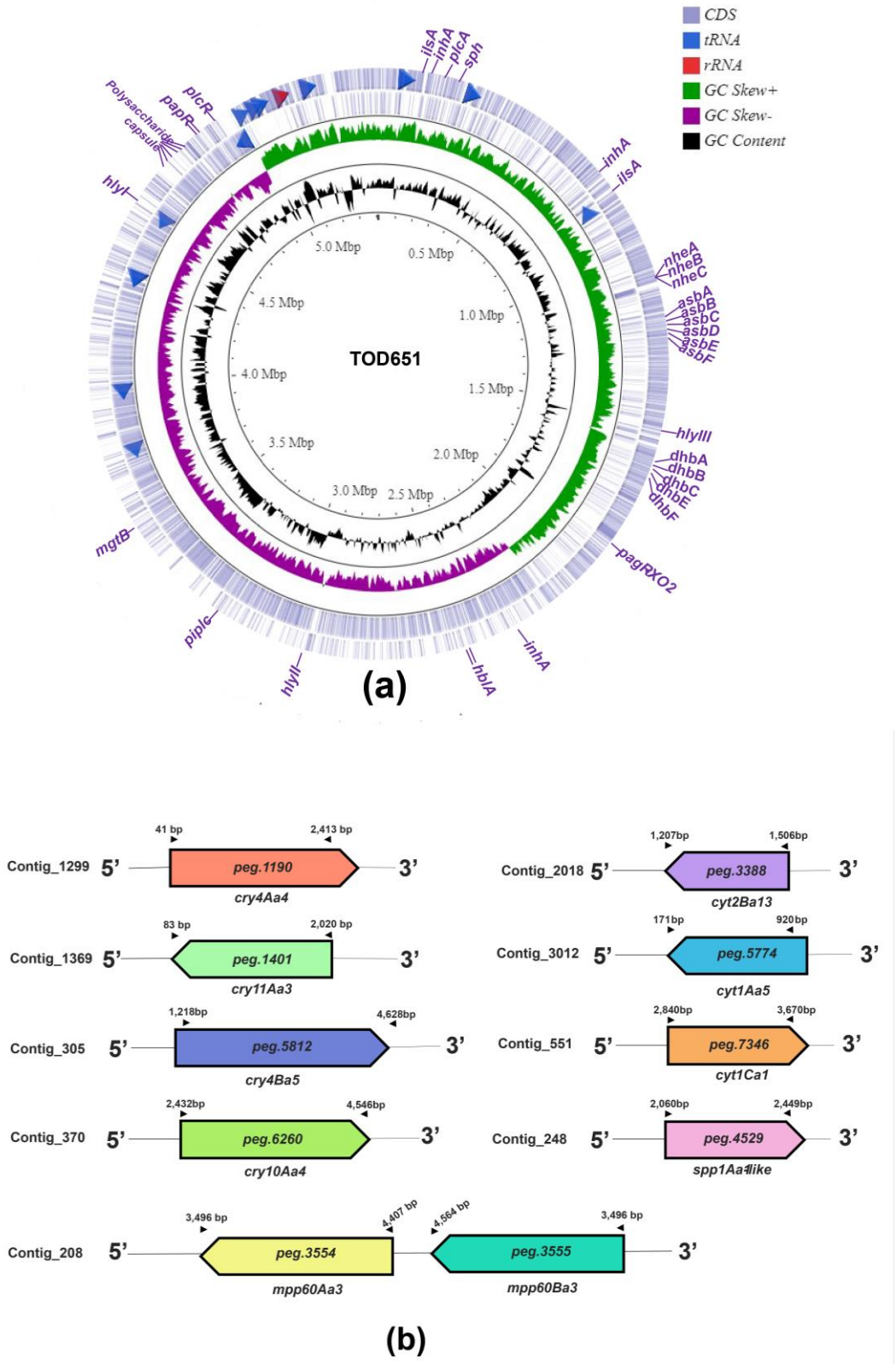

Figure 4. A graphical representation of the virulence factors and pesticidal protein-like genes found in the draft genome of *Bt* TOD651. **(a)** Distribution of virulence factor genes in draft chromosome. **(b)** Position of pesticidal protein-like genes in contigs unused in chromosome assembly.

A total of 10 CDS predicted in the *Bt* TOD651 genome were highly homologous to pesticidal proteins (Figure 4 b, Table 3). Four cry genes (*cry11Aa3*, *cry10Aa4*, *cry4Aa4* and *cry4Ba5*), three cyt (*cyt1Aa5*, *cyt1Ca1* and *cyt2Ba13*), two mpp (*mpp60Aa3* and *mpp60Ba3*) and one spp (*spp1Aa1*) genes were identified (Figure 4 b, Table 3).

Table 3. Identification of genes coding pesticidal proteins-like in the *Bt* TOD651.

Sequence	CDS predict	Length (aa)	Protein Homologous	Coverage (%)	Pairwise Identity (%)	E-value
Contig_1299	peg.1190	698	Cry4Aa4	59.07[1]/99.90[2]	99.43[1,2]	0.0
Contig_1369	peg.1401	645	Cry11Aa3	100.00[1,2]	100.00[1,2]	0.0
Contig_305	peg.5812	1161	Cry4Ba5	100.00[1,2]	100.00[1,2]	0.0
Contig_370	peg.6260	674	Cry10Aa4	97.19[1]/ 100.00[2]	100.00[1,2]	0.0
Contig_2018	peg.3388	99	Cyt2Ba13	40.24[1]/ 100.00[2]	100.00[1,2]	0.0
Contig_3012	peg.5774	262	Cyt1Aa5	100.00[1]	100.00[1]	0.0
Contig_551	peg.7346	291	Cyt1Ca1	51.43[1]/97.80[2]	98.90[1,2]	0.0
Contig_208	peg.3554	323	Mpp60Aa3	100.00[1,2]	100.00[1,2]	0.0
Contig_208	peg.3552	319	Mpp60Ba3	100.00[1,2]	100.00[1,2]	0.0
Contig_248	peg.4529	323	Spp1Aa1	58.70[1]	80.81[1]	0.0

[1]Btoxin_Digger
[2]Customized

3.5 Proteomics of spore-crystal mixture

The detected protein sequences were functionally classified into 10 GO terms related to cellular components, eight GO terms related to molecular functions, and 14 terms related to biological processes (Figure 5). In the cellular component groups, most proteins were related mainly to cell part and cell, the molecular function classification was represented by proteins with catalytic and binding activities; and in the biological process category, mostly proteins belonged to metabolic and cellular processes (Figure 5).

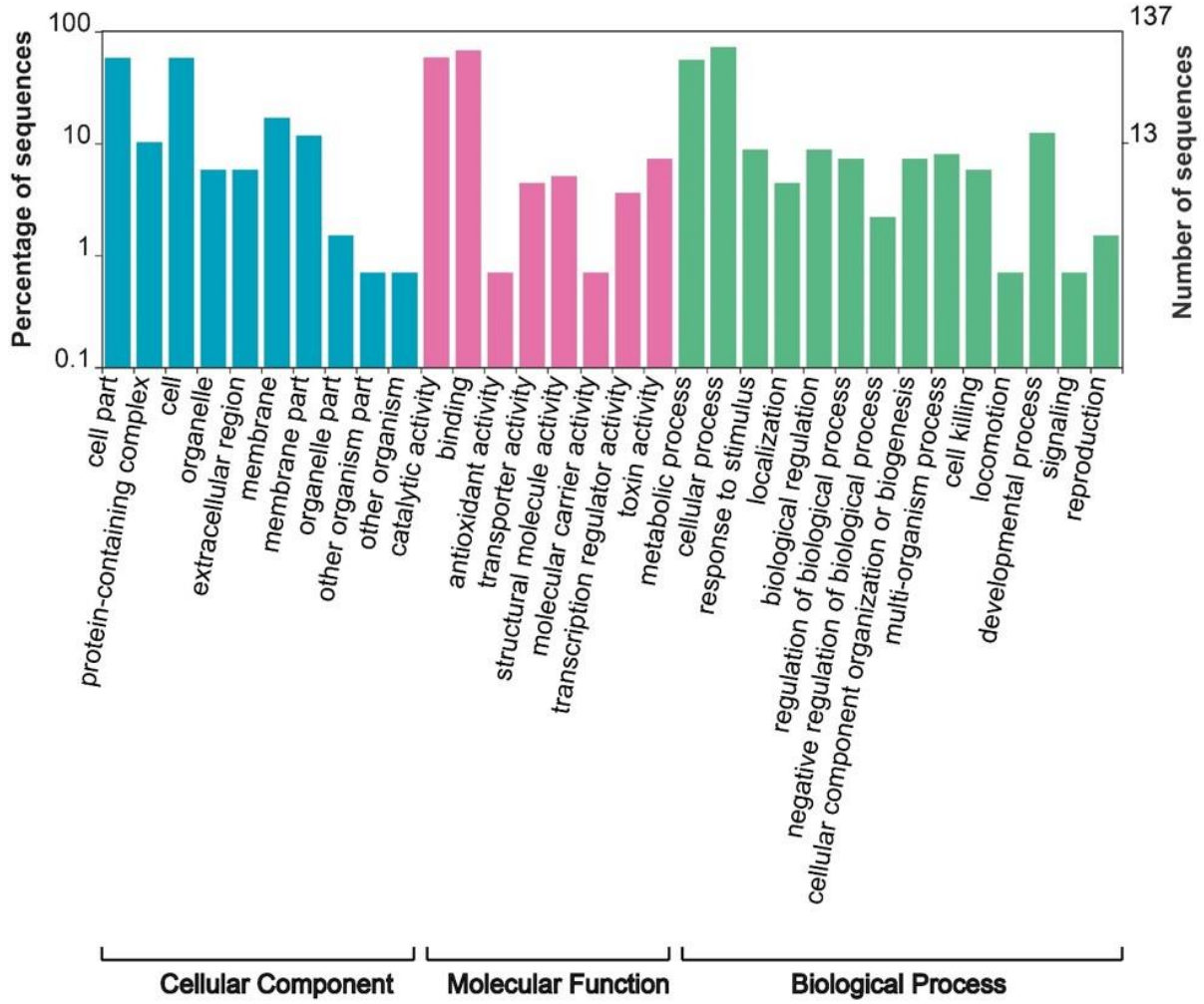

Figure 5. Functional annotation and classification of proteins identified in the spore-crystal mixture of TOD651.

A comparison of genomic and proteomic data was performed to identify predicted CDS that were expressed. A total of 43 CDS regions annotated in the genome were detected in the proteomic analysis with at least two peptides (Table 4). With respect to pesticidal proteins, only *mpp60Aa3* did not show a unique peptide, and therefore its expression was not confirmed. Thus, the expression of Cry11Aa3, Cry10Aa4, Cry4Aa4, Cry4Ba5, Cyt1Aa5, Cyt1Ca1, Cyt2Ba13, and Mpp60Ba3 were confirmed. Cry4Ba5 was the most abundant peptide discovered (60), followed by Cry4Aa4 (47), Cry11Aa3 (46), Mpp60Aa3 (29), and Cry10Aa4 (29) (Table 4, Supplemental Table S2). Among the cytolytic proteins, Cyt1Aa was the most abundant and showed 13 unique peptides, while Cyt2Ba13 and Cyt1Ca1 showed seven and five unique peptides, respectively (Table 4, Supplemental Table S2). Besides pesticidal proteins, other proteins were identified. Metallophosphoesterase was the most abundant protease and showed 25 unique peptides (Table 4). Three unique peptides were

from the virulence factor, extracellular neutral protease B. Proteins involved in sporulation (spore coat proteins, and exosporium protein), protein biosynthesis (chaperone protein, heat shock protein, translation elongation factor Tu, ribosomal proteins), and other functions (e.g., aminopeptidase, enolase, DNA-binding protein, RNA-binding proteins and Superoxide dismutase) were also identified (Table 4).

Table 4. Pesticidal and other proteins identified in the spore-crystal mixture of the *Bt* TOD651.

CDS id	Description[1]	Length (bp)	Peptide sequence (n°)	Unique Peptide (n°)[4]	Coverage[5]	Protein Score[6]	NSAF[7]
peg.5812	Cry4Ba5[2]	1136	62	60	0.5599	211.396	0.0492475
peg.1190	Cry4Aa4[2]	791	51	47	0.6587	184.808	0.0732234
peg.1401	Cry11Aa3[2]	645	46	46	0.6171	161.116	0.1816369
peg.3553	Mpp60Ba3[2]	303	29	29	0.8119	97.915	0.0868883
peg.6260	Cry10Aa4[2]	705	31	29	0.4057	104.199	0.031742
peg.3039	Metallophosphoesterase (Mppe)[3]	471	25	25	0.7113	91.843	0.0440183
peg.5774	Cyt1Aa5[2]	249	13	13	0.6345	50.49	0.1850302
peg.1543	Heat shock protein (GroEL)	544	9	9	0.1397	27.31	0.0060494
peg.5502	L-alanyl-gamma-D-glutamyl-L-diamino acid endopeptidase	325	9	9	0.4215	33.37	0.0131636
peg.3700	Spore coat protein (CotB)	174	8	8	0.546	28.416	0.0264785
peg.3439	Aminopeptidase	466	8	8	0.2833	27.635	0.0063558
peg.4239	Elongation factor Tu	320	8	8	0.3906	31.32	0.0133693
peg.8774	Enolase	431	7	7	0.2877	27.999	0.0068719
peg.2515	Spore coat protein (CotG)	186	7	7	0.1828	22.946	0.0212316
peg.3388	Cyt2Ba13[2]	99	7	7	0.6768	24.895	0.0465379
peg.7343	Cyt1Ca1[2]	85	5	5	0.2588	14.095	0.0309731
peg.7029	Dihydrolipoamide dehydrogenase of pyruvate dehydrogenase complex	470	4	4	0.0723	11.744	0.003501
peg.8531	Chaperone protein (DnaK)	611	4	4	0.0917	12.256	0.0021544
peg.5180	Spore coat protein (GerQ)	139	4	4	0.295	12.928	0.0165728
peg.5234	LSU ribosomal protein L7p/L12p(P1/P2)	119	4	4	0.5462	13.206	0.0110618
peg.7670	NAD-dependent glyceraldehyde-3-phosphate dehydrogenase	334	4	4	0.2395	16.849	0.0049265
peg.7030	Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex	227	4	4	0.1674	8.968	0.0072487
peg.7476	Hypothetical protein	250	3	3	0.156	10.916	0.0052654
peg.2818	Fructose-bisphosphate aldolase class II	267	3	3	0.1723	9.314	0.0049302
peg.1753	N-acetylmuramoyl-L-alanine amidase	271	3	3	0.1144	9.547	0.0048574
peg.8176	Uncharacterized protein YmfJ	82	3	3	0.4634	8.95	0.0120399
peg.4909	N-acetylmuramoyl-L-alanine amidase	327	3	3	0.1437	8.267	0.0030192

peg.3986	Extracellular neutral protease B (NprB)	426	3	3	0.1244	11.317	0.0038626
peg.9127	SSU ribosomal protein S2p (SAe)	233	2	2	0.1159	5.189	0.0028248
peg.7282	Hypotetical protein	143	2	2	0.1678	4.502	0.0046026
peg.4363	Superoxide dismutase	203	2	2	0.1823	6.977	0.0048634
peg.782	Hypothetical protein	129	2	2	0.186	5.041	0.0051022
peg.4419	Cell division trigger factor	404	2	2	0.0965	7.55	0.0016292
peg.3911	DNA-binding protein (Hbsu)	90	2	2	0.3667	7.282	0.0073131
peg.7358	Spore coat protein (CotE)	180	2	2	0.2111	6.641	0.0036565
peg.969	Phage tail fiber protein	431	2	2	0.051	4.816	0.0015271
peg.3496	Hypotetical protein	108	2	2	0.1481	4.385	0.0060942
peg.9081	Uncharacterized protein BA5373	68	2	2	0.5	6.717	0.0096791
peg.8583	RNA-binding protein (Hfq)	74	2	2	0.3243	6.323	0.0088943
peg.3589	Spore coat protein of CotY/CotZ family	155	2	2	0.2	6.649	0.0063695
peg.1544	Heat shock protein (GroES)	94	2	2	0.2447	5.286	0.0070019
peg.5565	Tricarboxylate transport sensor protein (TctE)	92	2	2	0.3261	3.967	0.0071541
peg.8113	Exosporium protein K	118	2	2	0.1102	5.905	0.0083667

[1]Annotation based on RASTtk
[2]Classification based on Btoxin_Digger and/or Customized *Bt* database
[3]Descrition based on BLASTx
[4]The number of peptide sequences that are unique to protein.
[5]The percentage of the protein sequence covered by identified peptides.
[6]The sum of the ion scores of all peptides that were identified.
[7]Normalized Spectral Abundance Factor, calculated using the number of spectra divided by the protein length and then normalized over the total of spectral counts/length for all the proteins in the sample.

4 Discussion

Here, we demonstrated that a novel Neotropical *Bt* strain (TOD651), isolated from Brazilian soil samples, exhibited significant larvicidal activities against larvae of *A. aegypti* and *C. quinquefasciatus*. Such larvicidal activities were similarly potent to those recorded in *Bt* strains that are already commercialized. Besides the characterization of the potential of TOD651 to be integrated into biorational programs of mosquito control, we further explored the TOD651 genome and proteome and discovered that its pathogenicity may be derived from the expression of pesticidal proteins (e.g. Cry11Aa3, Cry10Aa4, Cry4Aa4, Cry4Ba5, Cyt1Aa5, Cyt2Ba13, and Mpp60Ba3), virulence factors (e.g., Metallophospoesterase and NprB) and coat proteins (e.g., CoatB, CoatE, CoatG, and CotY/CotZ family protein).

The protein profile and ultrastructure of the parasporal crystals of the TOD651 strain were observed to be similar to those of other anti-dipteran *Bt* strains, including the reference strain [44,45]. The SDS-PAGE 130 kDa protein band represents the Cry4 protein, the 70kDa band indicates the presence of Cry11/Cry10, and the 27 kDa band suggests the presence of the Cyt protein [4,8]. Besides, the spherical crystals observed by scanning electron microscopy are usually the crystal morphology found in this type of *Bt* strain [46].

The phylogenetic analysis based on the *gyrB* gene showed that *Bt* TOD651 was closely related to *Bti* strains. In addition, a nematocidal strain (*Bt* MYBT18246) [47], a type-strain (*Bt* ATCC 10792), and a strain toxic against lepidopteran insects (*Bt* serovar thuringiensis IS5056) [48] were also closely related to *Bt* TOD651.

The TOD651 strain's whole-genome analysis revealed the following genes: *cry11Aa3*, *cry10Aa4*, *cry4Aa4*, *cry4Ba5*, *cyt1Aa5*, *cyt1Ca1*, *cyt2Ba13*, *mpp60Aa3*, and *mpp60Ba3*. Other genomic studies revealed similar gene content in *Bt* strains with high mosquitocidal activity. For example, *Bt* AR23 has been described to harbor *cry10Aa4*, *cry11Aa3*, *cry4Ba5*, *cry4Aa4*, *cyt2Ba*, *cyt1Aa* and *cyt1Ca* [28], while *Bt* LLP29 harbors *cry4Aa4*, *cry10Aa4*, *cry11Aa4*, *cyt1Aa6*, *cyt2Ba1* and *cry22Aa* [29]. The *Bt* TOD651 genome also harbors different virulence factors genes, such as haemolysins, enterotoxins, proteases, and phospholipases. These virulence factor genes, conserved in the *Bacillus cereus* group, are responsible for the colonization and adaptation of *Bt* in insect hosts [49-51].

The proteomic analysis revealed the expression of Cry11Aa3, Cry10Aa4, Cry4Aa4, Cry4Ba5, Cyt1Aa5, Cyt2Ba13, and Mpp60Ba3 proteins. Thus, these proteins were responsible for the toxic activity of TOD651 against *A. aegypti* and *C. quinquefasciatus*. Stein et al. [63] detected the transcripts of *cyt1Ca* but did not find Cyt1Ca protein. Here, Cyt1Ca1protein was expressed. However, it may not be related with the toxicity of *Bt* TOD651, since neither mosquito larvicidal activity function has been reported for Cyt1Ca [64].

Bt TOD651 showed the lower CL_{50} for *A. aegypti*. When tested individually, mosquitocidal Cry proteins have different toxicity levels between *Aedes* and *Culex* genera. The Cry4Ba protein is highly toxic against *A. aegypti*, while the Cry4Aa is highly toxic for Culex mosquitoes [52,53]. Cry11Aa has been linked to high toxicity against both the *Aedes* and *Culex* genera [53], whereas Cry10Aa is toxic to *A. aegypti* insects [8]. Mpp60 proteins show moderate insecticidal activity against *C. quinquefasciatus* larvae [54]. Cyt1Aa and Cyt2Ba exhibit toxicity towards both *A. aegypti* and *C. quinquefasciatus* [55,56]. However, these proteins act synergistically, which is mainly attributed to the Cyt1Aa toxin that can increase the activity of Cry4Aa, Cry4Ba, Cry11Aa, or Cry10Aa [57].

Cry10Aa has been reported as the minor protein component of *Bti* crystals [6]. In the proteomic analysis of *Bti* 4Q2-72, Cry10Aa presented low abundance of peptides (Fu et al., 2008). Besides, Cyt2Aa is also considered a minority pesticidal protein produced by *Bti* [57]. In the proteomic analysis of the commercial *Bti* AM65-52, Cry10Aa and Cyt2Ba were not

expressed [19]. As well as the Cyt1Aa protein, the Cyt2Ba protein is highly synergistic with the Cry proteins, and hence their combinations prevent the emergence of resistance in the target insects [4]. In addition, Cry10Aa seems to contribute to the overall toxicity of *Bti* [4]. The Cry10Aa4 was one of the most abundant proteins in spore-crystal mixture of *Bt* TOD651 and the peptides number of Cyt2Ba13 was sufficient to not consider it as a trace-level protein.

Cyt1Aa has been reported as the major component of *Bti* AM65-52 crystals [19], while *Bti* 4Q2-72 expressed higher level of Cry11Aa [65]. In contrast, *Bt* TOD651 expressed mainly Cry4Ba5. Cyt1Aa play an important role in the synergism of *Bti* strains and may also contribute to delay the evolution of resistance to Cry proteins in low proportions [68].

Regarding to Mpp proteins, only Mpp60Ba3 expression was detected in this study. Despide the fact that the mpp60Aa3 and mpp60Ba3 genes are part of the same operon, neither of them depends on the other to be expressed [54]. SDS-PAGE gel analysis revealed no detectable protein band for Mpp60Ba3. Sauka et al. [58] discovered an *mpp* homolog in a *B. toyonensis* strain's genome but no protein bands by SDS-PAGE analysis, implying that these proteins are secreted and present remnant fractions. The expression of Mpp60Aa and Mpp60Ba proteins have been detected in *Bti* AM65-52 [39] and *Bt jegathesan* [54] in low abundance. In contrast, Mpp60Ba3 represented high proportion in *Bt* TOD651, suggesting that it plays an important role in the toxicity of this bacteria.

Metalloproteinases have been described as virulence factors involved in *Bt* pathogenesis increasing the toxicity of pesticidal proteins [8]. The immune inhibitor A (InhA) has been identified as the main metalloproteinase associated with pesticidal proteins in spore-mixture crystals of *Bt* strains [25,66]. Interestingly, in this study, the high protein homology to Metallophosphoesterase was the most abundant protease in *Bt* TOD651. The gene of Metallophosphoesterase family has been identified in *B. cereus* genome [67]. This metalloprotease should play important role in the stress resistance of bacteria to adapt to the environment/host [67]. Neutral protease B (NprB) (also named NprA and Npr99), also present in the spore-crystal mixture of *Bt* TOD651, has been associated with the virulence of *B. cereus*, degrading host tissues and increasing tissue permeability to the pathogen [60].

Other proteins that may contribute to *Bt* TOD651 toxicity were detected in the spore-crystal mixture, such as spore coat proteins. It has been demonstrated that spores in the spore-crystal mixture play a major role in *Bt* toxicity, not only due to septicemia from spore germination and outgrowth but also due to synergy between the spore coat protein and crystal

protein [59]. Heat-shock proteins and the elongation factor Tu were also detected, which are necessary for the formation of crystals in *Bt* strains [61,62].

The genomic and proteomic analysis of the *Bt* TOD651 and other *Bt* strains can provide insights into the genetic makeup and protein expression of the bacteria. This information can help identify the genes responsible for the bacterium's insecticidal properties, which can be used for developing new, more effective insecticides. Additionally, genomic and proteomic analysis can provide information on the evolutionary history and diversity of different *Bt* strains, which can inform their classification and aid in the development of new strains with desired traits.

The *Bt* TOD651 can be used as an alternative for *A. aegypti* and *C. quinquefasciatus* control, and the combined genomic and proteomic analysis revealed the proteins directly related to their toxicity. Besides showed the pesticidal proteins proportions different of commercial strain, *Bt* TOD651 exhibit more varied of protein content that can potentialize delay the evolution of resistance. Furthermore, we detected other proteins can also contribute to *Bt* TOD651 pathogenicity.

Supplementary Materials: Table S1. Virulence factors identified in the chromosome draft sequence of *Bt* TOD651 from the VFDB database; Table S2: Peptides identified to pesticidal proteins classified as unique (true) or non-unique (false).

5 References

1. Ramos-Nino M. E., Fitzpatrick, D. M., Eckstrom, K. M., Tighe, S., Hattaway, L. M., Hsueh, A. N., Cheetham, S. Metagenomic analysis of *Aedes aegypti* and *Culex quinquefasciatus* mosquitoes from Grenada, West Indies. *Plos one*. **2020**, *15*, e0231047.

2. Boukedi, H., Hman, M., Khedher, S. B., Tounsi, S., Abdelkefi-Mesrati, L. Promising active bioinsecticides produced by *Bacillus thuringiensis* strain BLB427. *J. Adv. Res. Rev.* **2020,** *8*, 026-035.

3. AL-FAR, I. M. *Bacillus thuringiensis* and its pesticidal crystal proteins. *Int. J. Fauna Biol*. **2020**, *3*, p. 157-162.

4. Ben-Dov, E. *Bacillus thuringiensis* subsp. *israelensis* and its dipteran-specific toxins. *Toxins.* **2014**, 1222-1243.

5. Lutinski, J. A., Kuczmainski, A. G., de Quadros, S., Busato, M. A., Weirich, C. M. M., Malgueiro, A., Garcia, F. R. M. *Bacillus thuringiensis* var. *israelensis* como alternativa para o controle populacional de *Aedes aegypti* (Linnaeus, 1762) (Diptera: Culicidae). *CeN.* **2017**, *39*, 211-220.

6. Berry, C., O'Neil, S., Ben-Dov, E., Jones, A. F., Murphy, L., Quail, M. A., ... & Parkhill, J. Complete sequence and organization of pBtoxis, the toxin-coding plasmid of *Bacillus thuringiensis* subsp. *Israelensis*. *Appl. Environ. Microbiol*. **2002**, *68*, 5082-5095.

7. Gray, E. W.; Fusco, R. Microbial control of black flies (Diptera: Simuliidae) with *Bacillus thuringiensis* subsp. *israelensis*. In: Microbial Control of Insect and Mite Pests. *Academic Press.* p. 2017367-377.

8. Valtierra-de-Luis D, Villanueva M, Berry C, Caballero P. Potential for *Bacillus thuringiensis* and Other Bacterial Toxins as Biological Control Agents to Combat Dipteran Pests of Medical and Agronomic Importance. *Toxins.* **2020**, 773.

9. Viana, J. L., Soares-da-Silva, J., Vieira-Neta, M. R. A., Tadei, W. P., Oliveira, C. D., Abdalla, F. C., ... & Pinheiro, V. C. S. Isolates of *Bacillus thuringiensis* from Maranhão biomes with potential insecticidal action against *Aedes aegypti* larvae (Diptera, Culicidae). *Braz. J. Biol.* **2020**, *81*, 114-124.

10. Zhou, Y., Wu, Z., Zhang, J., Wan, Y., Jin, W., Li, Y., & Fang, X. (2020). *Bacillus thuringiensis* novel toxin Epp is toxic to mosquitoes and *Prodenia litura* larvae. *Braz. J. Microbiol*. **2020**, *51*, 437-445.

11. Wu, J., Wei, L., He, J., Fu, K., Li, X., Jia, L., ... & Zhang, W. Characterization of a novel *Bacillus thuringiensis* toxin active against *Aedes aegypti* larvae. *Acta Trop.* **2021**, *223*, 106088.

12. Roy, M., Chatterjee, S., & Dangar, T. K. (2021). Characterization and mosquitocidal potency of a *Bacillus thuringiensis* strain of rice field soil of Burdwan, West Bengal, India. *Microb. Pathog.* **2021**, *158*, 105093.

13. Day M., Ibrahim M., Dyer D., Bulla L. Genome Sequence of *Bacillus thuringiensis* subsp. kurstaki Strain HD-1. *Genome Announc*. **2014**, *2*, e00613-00614.

14. Jeong H., Choi S-K., Park S-H. Genome Sequences of *Bacillus thuringiensis* Serovar *kurstaki* Strain BP865 and *B. thuringiensis*, Serovar *aizawai* Strain HD-133. *Genome Announc*. **2017**, 5:e01544-01516.

15. Sajid, M., Geng, C., LI, M., Wang, Y., Liu, H., Zheng, J., Sun, M. Whole-genome analysis of *Bacillus thuringiensis* revealing partial genes as a source of novel *cry* toxins. *Appl. Environ. Microbiol.* **2018**, *84*, e00277-18.

16. Lechuga, A., Lood, C., Salas, M., van Noort, V., Lavigne, R., & Redrejo-Rodríguez, M. Completed genomic sequence of *Bacillus thuringiensis* HER1410 reveals a Cry-containing chromosome, two megaplasmids, and an integrative plasmidial prophage. *G3: Genes*, *Genomes*, *Genet.* **2020**, *10*, 2927-2939.

17. Susič, N., Janežič, S., Rupnik, M., & Stare, B. G. Whole genome sequencing and comparative genomics of two nematicidal Bacillus strains reveals a wide range of possible virulence factors. *G3: Genes*, *Genomes*, *Genetic.* **2020**, *10*, 881-890.

18. Wu, D., He, J., Gong, Y., Chen, D., Zhu, X., Qiu, N., ... & Yu, Z. Proteomic analysis reveals the strategies of *Bacillus thuringiensis* YBT-1520 for survival under long-term heat stress. *Proteomics*, **2011**, *11*, 2580-2591.

19. Caballero, J., Jiménez-Moreno, N., Orera, I., Williams, T., Fernández, A. B., Villanueva, M., ... & Ancín-Azpilicueta, C. Unraveling the composition of insecticidal crystal proteins in *Bacillus thuringiensis*: A proteomics approach. *Appl. Environ. Microbiol*., **2020,** *86*, e00476-20.

20. Rang, J., He, H., Wang, T., Ding, X., Zuo, M., Quan, M., ... & Xia, L. Comparative analysis of genomics and proteomics in *Bacillus thuringiensis* 4.0718. *PLoS One*, **2015**, *10*(3), e0119065.

21. Alves, G. B., Melo, F. L., Oliveira, E. E., Haddi, K., Costa, L., Dias, M. L., ... & Aguiar, R. W. Comparative genomic analysis and mosquito larvicidal activity of four *Bacillus thuringiensis* subsp. *israelensis* strains. Sci. Rep. **2020**, *10*, 1-12.

22. Dankocsik C, Donovan WP, Jany CS. Activation of a cryptic crystal protein gene of *Bacillus thuringiensis* subspecies kurstaki by gene fusion and determination of the crystal protein insecticidal specificity. *Mol Microbiol.* **1990**, *4*, 2087–2094.

23. Quan, M., Xie, J., Liu, X., Li, Y., Rang, J., Zhang, T., ... & Ding, X. Comparative analysis of genomics and proteomics in the new isolated *Bacillus thuringiensis* X022 revealed the metabolic regulation mechanism of carbon flux following Cu^{2+} treatment. *Front. microbiol.*, **2016**, *7*, 792.

24. Gomis-Cebolla J, Scaramal Ricietto AP, Ferré J. A genomic and proteomic approach to identify and quantify the expressed *Bacillus thuringiensis* proteins in the supernatant and parasporal crystal. *Toxins.* **2018**, *10*, 193.

25. Khorramnejad A, Gomis-Cebolla J, Talaei-Hassanlouei R, Bel Y, Escriche B (2020) Genomics and proteomics analyses revealed novel candidate pesticidal proteins in a lepidopteran-toxic *Bacillus thuringiensis* strain. *Toxins.* **2020**, *12*, 673.

26. Baragamaarachchi RY, Samarasekera JKRR, Weerasena OVDSJ, Lamour K, Jurat-Fuentes JL. Identification of a native *Bacillus thuringiensis* strain from Sri Lanka active against Dipel-resistant *Plutella xylostella. PeerJ.* **2019**, *7*, e7535.

27. Doggett, N. A., Stubben, C. J., Chertkov, O., Bruce, D. C., Detter, J. C., Johnson, S. L., & Han, C. S. Complete genome sequence of *Bacillus thuringiensis* serovar israelensis strain HD-789. *Genome announc*. **2013**, *1*, e01023-13.

28. Fayad, N., Patiño-Navarrete, R., Kambris, Z., Antoun, M., Osta, M., Chopineau, J., ... & Kallassy Awad, M. Characterization and whole genome sequencing of AR23, a highly toxic *Bacillus thuringiensis* strain isolated from Lebanese soil. *Curr. Microbiol.* **2019**, *76*, 1503-1511.

29. Ma, W., Chen, H., Jiang, X., Wang, J., Gelbič, I., Guan, X., & Zhang, L. Whole genome sequence analysis of the mosquitocidal *Bacillus thuringiensis* LLP29. *Arch. Microbiol.* **2020**, *202*, 1693-1700.

30. Monnerat, R. G., Batista, A. C., de Medeiros, P. T., Martins, E. S., Melatti, V. M., Praça, L. B., ... & Berry, C. Screening of Brazilian *Bacillus thuringiensis* isolates active against *Spodoptera frugiperda*, *Plutella xylostella* and *Anticarsia gemmatalis*. *Biol. Control*, **2007**, *41*, 291-295.

31. Mounsef, J. R., Salameh, D., kallassy Awad, M., Chamy, L., Brandam, C., & Lteif, R. A simple method for the separation of *Bacillus thuringiensis* spores and crystals. *J. Microbiol. Methods.* **2014**, *107*, 147-149.

32. Aguiar, R. W. S., dos Santos, S. F., da Silva Morgado, F., Ascencio, S. D., de Mendonça Lopes, M., Viana, K. F., ... & Ribeiro, B. M. Insecticidal and repellent activity of Siparuna guianensis Aubl. (Negramina) against *Aedes aegypti* and *Culex quinquefasciatus. PloS one*. **2015,** *10*, e0116765.

33. McLaughlin, R. E., Dulmage, H. T., Alls, R., Couch, T. L., Dame, D. A., Hall, I. M., ... & Versoi, P. L. (). US standard bioassay for the potency assessment of *Bacillus thuringiensis* serotype H-14 against mosquito larvae. *Bulletin of the ESA.* **1984**, *30* 26-29.

34. SAS Institute. SAS/STAT User's Guide., Cary, NC, USA, **2008**.

35. Andrews S. FastQC: a quality-control tool for high-throughput sequence. **2015.** https://www.bioinformatics.babraham.ac.uk/projects/fastqc. Accessed October 2021

36. Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., ... & Drummond, A. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. **2012**, *28*, 1647-1649.

37. Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., ... & Pevzner, P. A. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **2012**, *19*, 455-477.

38. Lischer, H. E., & Shimizu, K. K. Reference-guided de novo assembly approach improves genome reconstruction for related species. *BMC bioinform*. **2017**, *18*, 1-12.

39. Liu, B., Zheng, D., Jin, Q., Chen, L., & Yang, J. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. Nucleic Acids Res. **2019**, *47*, D687-D692.

40. Liu H, Zheng J, Bo D, Yu Y, Ye W, Peng D, Sun M. BtToxin_Digger: a comprehensive and high-throughput pipeline for mining toxin protein genes from *Bacillus thuringiensis. Bioinformatics.* **2021**, *38*, 250-251.

41. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol*. **2018**, *35*, 1547.

42. Santos, M. D., Lima, D. B., Fischer, J. S., Clasen, M. A., Kurt, L. U., Camillo-Andrade, A. C., ... & Carvalho, P. C. Simple, efficient and thorough shotgun proteomic analysis with PatternLab V. *Nat. Protoc*. **2022**, 1-26.

43. Ye, J., Zhang, Y., Cui, H., Liu, J., Wu, Y., Cheng, Y., ... & Shi, C. WEGO 2.0: a web tool for analyzing and plotting GO annotations, 2018 update. *Nucleic Acids Res.* **2018**, *46*, W71-W75.

44. Elleuch, J., Jaoua, S., Darriet, F., Chandre, F., Tounsi, S., & Zghal, R. Z. Cry4Ba and Cyt1Aa proteins from *Bacillus thuringiensis israelensis*: Interactions and toxicity mechanism against Aedes aegypti. *Toxicon.* **2015**, *104*, 83-90.

45. El-Kersh, T. A., Ahmed, A. M., Al-Sheikh, Y. A., Tripet, F., Ibrahim, M. S., & Metwalli, A. A. Isolation and characterization of native *Bacillus thuringiensis* strains from Saudi Arabia with enhanced larvicidal toxicity against the mosquito vector *Anopheles gambiae* (sl). *Parasites & vectors.* **2016**, *9*, 1-14.

46. Nair, K., Al-Thani, R., Al-Thani, D., Al-Yafei, F., Ahmed, T., & Jaoua, S. Diversity of *Bacillus thuringiensis* strains from Qatar as shown by crystal morphology, δ-endotoxins and cry gene content. *Front. microbiol*., **2018**, *9*, 708.

47. Hollensteiner, J., Poehlein, A., Spröer, C., Bunk, B., Sheppard, A. E., Rosentstiel, P., ... & Liesegang, H. Complete genome sequence of the nematicidal *Bacillus thuringiensis* MYBT18246. *Stand Genomic Sci.* **2017**, *12*, 1-10.

48. Swiecicka, I., Bideshi, D. K., & Federici, B. A. Novel isolate of *Bacillus thuringiensis* subsp. *thuringiensis* that produces a quasicuboidal crystal of Cry1Ab21 toxic to larvae of Trichoplusia ni. *Appl. Environ. Microbiol.* **2008**, *74*, 923-930.

49. Vilas-Bôas, G. T., Alvarez, R. C., dos Santos, C. A., & Vilas-Boas, L. A. Fatores de virulência de *Bacillus thuringiensis*: o que existe além das proteínas Cry. *EntomoBrasilis.* **2012**, *5*, 1-10.

50. Guillemet, E., Cadot, C., Tran, S. L., Guinebretiere, M. H., Lereclus, D., & Ramarao, N. The InhA metalloproteases of *Bacillus cereus* contribute concomitantly to virulence. *J. Bacteriol*. **2010**, *192*, 286-294.

51. Pohare, M. B., Wagh, S. G., & Udayasuriyan, V. *Bacillus thuringiensis* as potential biocontrol agent for sustainable agriculture. In *Current Trends in Microbial Biotechnology for Sustainable Agriculture*, Springer, Singapore. **2021**. 439-468

52. Otieno-Ayayo, Z. N., Zaritsky, A., Wirth, M. C., Manasherob, R., Khasdan, V., Cahan, R., & Ben-Dov, E. Variations in the mosquito larvicidal activities of toxins from *Bacillus thuringiensis* ssp. *israelensis*. *Environ. Microbiol*. **2008**, *10*, 2191–2199.

53. Abdullah, M. A. F., Alzate, O., Mohammad, M., McNall, R. J., Adang, M. J., & Dean, D. H. Introduction of Culex toxicity into *Bacillus thuringiensis* Cry4Ba by protein engineering. *AppL. Environ. Microb.* **2003**, *69*, 5343–5353.

54. Sun, Y., Zhao, Q., Xia, L., Ding, X., Hu, Q., Federici, B. A., & Park, H. W. Identification and characterization of three previously undescribed crystal proteins from *Bacillus thuringiensis* subsp. *jegathesan*. *Appl. Environ. Microbiol*. **2013**, *79*, 3364-3370.

55. Juárez-Pérez, V.; Guerchicoff, A.; Rubinstein, C.; Delecluse, A. Characterization of Cyt2Bc toxin from *Bacillus thuringiensis* subsp. *medellin*. *Appl. Environ. Microbiol.* **2002**, *68*, 1228–1231.

56. Torres-Quintero, M.C.; Gómez, I.; Pacheco, S.; Sánchez, J.; Flores, H.; Osuna, J.; Mendoza, G.; Soberón, M.; Bravo, A. Engineering *Bacillus thuringiensis* Cyt1Aa toxin specificity from dipteran to lepidopteran toxicity. *Sci. Rep.* **2018**, *8*, 1–12.

57. Cantón, P.E.; Reyes, E.Z.; Ruiz De Escudero, I.; Bravo, A.; Soberon, M. Binding of *Bacillus thuringiensis* subsp. *israelensis* Cry4Ba to Cyt1Aa has an important role in synergism. *Peptides*, **2011**, *32*, 595–600.

58. Sauka, D. H., Peralta, C., Pérez, M. P., Onco, M. I., Fiodor, A., Caballero, J., ... & Palma, L. *Bacillus toyonensis* biovar *Thuringiensis*: A novel entomopathogen with insecticidal activity against lepidopteran and coleopteran pests. *Biol. Control.* **2022,** *167*, 104838.

59. Johnson, D. E., Oppert, B., & McGaughey, W. H. Spore coat protein synergizes *Bacillus thuringiensis* crystal toxicity for the Indianmeal moth (Plodia interpunctella). *Curr. Microbiol*. **1998**, *36*, 278-282.

60. Chung, M. C., Popova, T. G., Millis, B. A., Mukherjee, D. V., Zhou, W., Liotta, L. A., ... & Popov, S. G. Secreted neutral metalloproteases of *Bacillus anthracis* as candidate pathogenic factors. *J. Biol. Chem*, **2006**, *281*, 31408-31418.

61. Ding, X., Huang, J., Xia, L., Li, X., Yuan, C., & Dan, S. A proteomic analysis approach to study insecticidal crystal proteins from different strains of *Bacillus thuringiensis*. *Biocontrol Sci. Technol*. **2009**, *19*, 289-299.

62. Xie, J., Peng, J., Yi, Z., Zhao, X., Li, S., Zhang, T., ... & Ding, X. Role of hsp20 in the Production of Spores and Insecticidal Crystal Proteins in *Bacillus thuringiensis*. *Front. Microbiol*. **2019**, *10*, 2059.

63. Stein, C.; Jones, G.W.; Chalmers, T.; Berry, C. Transcriptional analysis of the toxin-coding plasmid pBtoxis from *Bacillus thuringiensis* subsp. *israelensis*. *Appl. Environ. Microbiol.* **2006**, 72, 1771–1776.

64. Cohen, S.; Albeck, S.; Ben-Dov, E.; Cahan, R.; Firer, M.; Zaritsky, A.; Dym, O. Cyt1Aa toxin: Crystal structure reveals implications for its membrane-perforating function. *J. Mol. Biol*. **2011**, 413, 804–814.

65. Fu, Z., Sun, Y., Xia, L., Ding, X., Mo, X., Li, X., ... & Zhang, Y. Assessment of protoxin composition of *Bacillus thuringiensis* strains by use of polyacrylamide gel block and mass spectrometry. *Appl. Environ. Microbiol*. **2008**, *79*, 875-880.

66. Banik, A., Chattopadhyay, A., Ganguly, S., & Mukhopadhyay, S. K. Characterization of a tea pest specific *Bacillus thuringiensis* and identification of its toxin by MALDI-TOF mass spectrometry. *Ind. Crops Prod*. **2019**, *137*, 549-556.

67. Zhang, J., Wang, H., Xie, T., Huang, Q., Xiong, X., Liu, Q., & Wang, G. The YmdB protein regulates biofilm formation dependent on the repressor SinR in *Bacillus cereus* 0–9. *World J. Microbiol. Biotechnol.* **2020**, *36*, 165.

68. Wirth, M. C., Park, H. W., Walton, W. E., & Federici, B. A. Cyt1A of *Bacillus thuringiensis* delays evolution of resistance to Cry11A in the mosquito *Culex quinquefasciatus*. *Appl. Environ. Microbiol*. **2005**, *71*, 185-189.

69. Valtierra-de-Luis, D., Villanueva, M., Lai, L., Williams, T., & Caballero, P. Potential of Cry10Aa and Cyt2Ba, two minority δ-endotoxins produced by *Bacillus thuringiensis* ser. israelensis, for the control of *Aedes aegypti* larvae. *Toxins*. **2020**, 12, 355.

CHAPTER III: Genomic analysis of a Neotropical *Bacillus thuringiensis* strain (UFT038) and its insecticidal potential against lepidopteran pests

Submitted to: Journal of Asia-Pacific Entomology (Annex C)

Abstract

Bacillus thuringiensis (*Bt*) is known for its Cry and Vip3A pesticidal proteins with high selectivity to target pests. Here, we assessed the potential of a novel Neotropical *Bt* strain (UFT038) against six lepidopteran pests, including two Cry-resistant populations of fall armyworm, *Spodoptera frugiperda*. In toxicological bioassays, UFT038 killed and inhibited the neonate growth in a concentration-dependent manner. Relative susceptibility to UFT038 was highest in *Chrysodeixis includens*, followed by *Helicoverpa armigera*, *H. zea*, *S. frugiperda* (susceptible population), *S. eridania*, and lowest in *S. cosmioides*. Compared to a commercial *Bt* serovar *kurstaki* (HD-1) formulation, UFT038 was 28-fold more toxic to *S. cosmioides*, while HD-1 was 70-fold more toxic than UFT038 to *C. includens*. Compared to HD-1, UFT038 was slightly less active against Cry1F and Cry1A.105+Cry2Ab resistant *S. frugiperda*. We sequenced and analyzed the UFT038 genome and identified virulence factor genes (*InhA1*, *plcA*, *piplc*, *sph*, and *chi1-2*) and toxins (*alo*, *cytK*, *hlyIII*, *hblA-D*, and *nheA*-C) beyond the insecticidal protein (*cry1Aa8*, *cry1Ac11*, *cry1Ia44*, *cry2Aa9*, *cry2Ab35*, and *vip3Af5)* genes. Collectively, our findings reveal the potential of the UFT038 strain as a source of pesticidal factors and genes against lepidopteran pests, including *S. cosmioides* and *S. frugiperda*.

Keywords: *Bacillus thuringiensis*; Lepidoptera; *Bt* resistance; genome sequencing, Cry proteins.

1 Introduction

Global food demand is growing with the increasing human population, imposing the need for increased agricultural production for sustainable food security. Grain crops such as soybean (*Glycine max* (L.) Merr.) and *maize* (Zea mays L.) are of great relevance for global food production. These crops, however, are attacked by several lepidopteran pests. Such lepidopteran-mediated damages cause high production and economic losses in agricultural crops worldwide (Sree and Varma, 2015), including in Brazilian fields (Freitas et al., 2019; Cordeiro et al., 2020).

Besides synthetic insecticides, the control of lepidopteran pests has relied chiefly on microbial-based agents such as *Bacillus thuringiensis* (*Bt*) and its insecticidal Cry and Vip proteins (Marrone, 2019). In addition to these insecticidal proteins, other virulence factors in the *Bt* chromosome make *Bt* a potential multi-target control tool (Deng et al., 2014; Zheng et al., 2017; Méric et al., 2018). In the last two decades, lepidopteran pest control has also been provided by transgenic crops transformed to produce Cry and Vip3A proteins (Romeis et al., 2019). However, continuous expression of *cry* and *vip3A* genes in *Bt* cultivars, their wide adoption, and lack of compliance with resistance management tools have resulted in the evolution of pest resistance to *Bt* proteins in transgenic crops (Tabashnik et al., 2017).

The increasing frequency of pest resistance to *Bt* proteins highlights the need to discover novel *Bt* isolates with potential for use in agriculture (Boonmee et al., 2019; Djenane et al., 2020; Baragamaarachchi et al., 2019). Efforts for biological and molecular characterization of novel *Bt* strains have greatly benefited from whole-genome sequencing to discover isolates with new insecticidal proteins and virulence factors (Zheng et al., 2017; Baragamaarachchi et al., 2019; Gomis-Cebolla et al., 2018; Zhu et al., 2015).

Here, we assessed the insecticidal activity of a novel Neotropical *B. thuringiensis* strain (UFT038) against different lepidopteran pest species (*Spodoptera frugiperda; S. eridania; S. cosmioides; Helicoverpa armigera*; *H. zea*; and *Chrysodeixis includens*), including Cry-resistant *S. frugiperda* strains. We further sequenced and analyzed the UFT038 genome, which allowed us to identify its gene content, including those related to virulence factors, and to conduct a phylogenetic analysis related to other *B. thuringiensis* strains.

2 Material and Methods

2.1 Bacterial strains

The UFT038 strain was isolated from soil samples collected in Tocantins state (Brazil) (11º43'45" S; 49º04'07" W). A commercial formulation based on *Bt* serovar *kurstaki* (*Btk*) HD-1 (Dipel WP 32 g/kg, 16,000 UI/mg, 25 billion viable spores/g) was purchased from Sumitomo Chemical Brasil Indústria Química S.A. (Barueri, São Paulo, Brazil).

The UFT038 strain was cultivated at 30°C for 24 h on Nutrient Yeast Extract Salt Medium (NYSM) agar medium. The HD-1 strain was prepared according to the manufacturer's recommendations, using distilled water as a diluent. Aliquots were transferred to Erlenmeyer's flasks (500 mL) containing Casein Hydrolysate and Yeast Extract (CCY) medium (Stewart et al., 1981) and incubated at 30°C, 200 rpm, for 72 h to produce spores and crystals (Azzouz et al., 2014; Zouari et al., 2002). The spores-crystals samples were centrifuged at 9700 × *g* for 10 min. The pellets were washed in 10 mL of sterile water, and this process was repeated two more times. The final pellet was resuspended in 5 mL of a 0.85% KCl solution and stored at 4°C. Spores and crystals in each solution were estimated under phase-contrast microscopy using a Neubauer chamber to prepare the active suspension used in bioassays. Spores-crystal suspensions of HD-1 were prepared using the same procedure.

2.2 Transmission electron microscopy

After cultivation in NYSM agar medium at 30ºC for 72 h, a loop of the colony sample was collected and diluted in sterile water. A volume of 100 µL was deposited over metallic supports and dried for 24 h at 37°C. Supports were then covered with gold for 180 s using an Emitech apparatus (model K550) and observed in a Zeiss scanning electron microscope (model DSM 962) at 10 or 20 Kv.

2.3 Insects and concentration-response bioassays

Six lepidopteran species and two Cry-resistant *S. frugiperda* populations were used. *Helicoverpa armigera*, *H. zea*, and *C. includens* eggs were purchased from Pragas.com Insumos Biológicos (http://www.pragas.com.vc/) (Piracicaba, São Paulo, Brazil). We used three populations of *S. frugiperda*, a standard susceptible laboratory population (S_Bt), one resistant to Cry1F (R_Cry1F) described elsewhere (Santos-Amaya et al., 2017a,b), and another strain resistant to Cry1A105+Cry2Ab (R_Cry1+2Ab) described previously (Santos-

Amaya et al., 2015). Populations of *S. cosmioides* and *S. eridania* previously described (Rabelo et al., 2020a,b) were also included in the bioassays. The insects were reared on an artificial diet based on common bean protein (*S. frugiperda*) (Kasten et al., 1978) or soybean derivatives and other ingredients (rest of species) (Greene et al., 1976). After three days, larvae were transferred to plastic trays containing one larva per well and reared until pupation. The pupae were sexed and transferred to Petri dishes containing a wet cotton swab for adequate moisture. Upon emergence, adults were placed in polyvinylchloride cylindrical cages lined with bond paper and fed 10% glucose and 1% ascorbic acid solution. The rearing was conducted at 27 °C ± 2 °C and 70 ± 10% RH and 10:14 (L: D) h photoperiod.

Diet-overlay bioassays were conducted by exposing neonates (< 24 h after hatching) in 128-well trays (CD International, Pitman, NJ, USA). As a positive comparator for toxicity, we used the commercial *Btk* HD-1 strain (Sumitomo Chemical Brasil Indústria Química S.A., Barueri, São Paulo, Brazil). We prepared seven spore-crystal concentrations diluting the original stocks in 0.1% Triton X-100, which was also used as a diluent in controls. One milliliter of rearing diet was placed per well, and after solidification, 30 µL of experimental or control solution was applied to the surface. After drying for 30–60 min at room temperature, a single neonate (< 24h hatching) was transferred to each well using a fine hair brush. The wells were sealed with lids (CD International, Pitman, NJ) and maintained at 27 ± 2ºC, 70 ± 10 % RH, 10:14 (L: D) h photoperiod. After seven days, mortality and mean larval weight were recorded. Larvae that did not grow beyond the first instar and weighed <0.1 mg were considered dead. Each treatment had four replications, with 16 larvae totaling 64 larvae tested per treatment. The data were analyzed using probit regression ($P > 0.05$) (Finney, 1971) in PoloPlus with adjustment for natural mortality when needed (Robertson et al., 2017). The susceptibility parameters estimated were the median lethal concentration (LC_{50}) with their respective 95% confidence limits (95% CL) and the response curve's slope to estimate the phenotypic variance in tolerance to *Bt*. Larval weights were transformed to the percentage of growth inhibition relative to the control larvae. This procedure allowed obtaining the effective concentration inhibiting 50% of larval growth (EC_{50}).

A series of lethal and effective concentration ratios were calculated in PoloPlus: i) the resistance ratio (ResR) of the insect populations to each *Bt* strain (UFT038 and HD-1); ii) the toxicity ratio (ToxR) between the two *Bt* strains for a given insect species or population, and iii) the cross-resistance ratio (CRR) of the Cry-resistant populations for each *Bt* strain. These ratios allow ranking the susceptibility of target pest species to *Bt* and determining their

relative potency and the magnitude of the change in susceptibility associated with Cry-Bt resistance. The above ratios and corresponding 95% confidence intervals were calculated using the appropriate reference and denominator, including the most sensitive insect strain, the most potent *Bt* strain, or the standard susceptible laboratory population of *S. frugiperda*, depending on each case. We considered the ratios significantly different ($P < 0.05$) when their 95% confidence interval did not include the 1.0 value (Robertson et al., 2017) or the 95% confidence limits did not overlap.

2.4 Genome sequencing, assembly, and annotation

Genomic DNA was extracted and purified using the Wizard® Genomic DNA Purification Kit (Promega, Madison, WI, USA). DNA concentration and purity were measured using the NanoDrop™ 8000 (Thermo Fisher Scientific, Waltham, MA, USA). The UFT038 genome was sequenced using the Illumina MiSeq™ platform with a coverage of 426x. A total of 14,732,122 paired-end reads of a mean length of 75.9 bp with an insert size of 200 bp were generated. The read quality was analyzed using FastQC v0.11.3 (Andrews, 2015), then reads were trimmed using a minimum Phred quality score >20. The reads were then normalized using BBNorm v.36.92 (Bushnell et al. 2014) within Geneious v.10.2.6 (Kearse et al. 2012), with a minimum depth of 5 and a target coverage level of 40 (Gaafar et al., 2020). The *de novo* assembly was performed using short-reads SPAdes assembler v 3.10.0 (within Geneious v.10.2.6) with default parameters (Bankevich et al., 2012). For chromosome assembly, the contigs obtained were aligned and reordered using *Btk* HD-1 (NCBI accession number CP004870) as reference sequence using the Geneious Map to reference tool (medium sensitivity and minimum overlap identity of 85%). The PLSDB v 2020_06_29 tool (Galata et al., 2019) was used to investigate plasmid-like sequences from contigs, with a maximum distance value of 0.05 and *P*-value of 0.1 (Wuyts et al., 2019). Contig extension was performed until genome completion or no further extension using Geneious and the assembled genome was annotated using the NCBI Prokaryotic genome annotation pipeline. The putative pesticidal protein gene content was determined using BtToxin_Digger (Liu et al., 2020). Chromosome sequence was functionally annotated with COG categories using eggNOG-mapper v.2 (Huerta-Cepas et al., 2019; Cantalapiedra et al., 2021) and potential virulence factors were predicted using the bacterial virulence factor database (VFDB) (Liu et al., 2019). The genomic sequence of *Bt* UFT038 was deposited in the NCBI database under BioProject PRJNA819055 and Genbank accession numbers CP094396 to CP094407.

2.5 Multilocus Sequence Typing (MLST) and phylogenetic analysis

Multilocus sequence typing (MLST) analysis was performed using PubMLST against the *Bacillus cereus* typing database (https://pubmlst.org/bcereus/). UFT038 chromosome was used as a sequence query for analyzing the seven housekeeping genes (*glp*, *gmk*, *ilvD*, *pta*, *pur*, *pycA*, and *tpi*). The *gyrB* gene (DNA gyrase subunit B) was detected by blastn from the contigs list using a custom database (*gyrB* genes list of *Bacillus* sp.) in the Geneious. Then, *gyrB* gene sequence from strain UFT038 was aligned to *gyrB* genes of closely related *Bt* strains retrieved from GenBank. We aligned the sequences using ClustalW (within Geneious) and created a phylogenetic tree using MEGA 11 (Tamura et al., 2021). We used the neighbor-joining method and bootstrap percentages based on 1000 replications. Bootstrap values <90% were disregarded for branches in the phylogenetic tree.

3 Results

3.1 Morphological analysis and concentration-response bioassays

Electron microscopy imaging detected the production of bipyramidal and cuboid crystals in the UFT038 strain (Fig. 1).

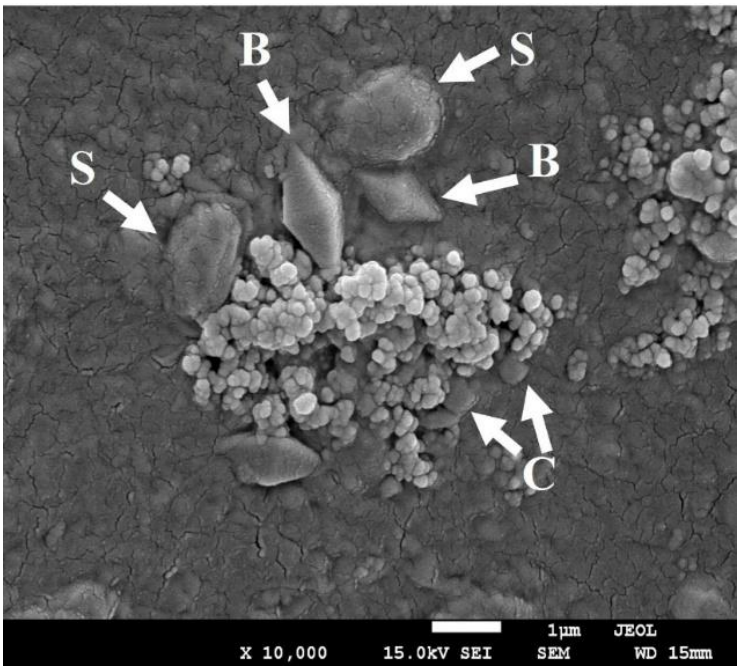

Fig.1 Ultrastructural characterization of spores and parasporal crystal proteins from the UFT038 strain. Arrows indicate spores (S), bipyramidal and cuboid crystals (B) observed. Spore and crystals were enlarged 2,000 (x) times.

The UFT038 strain showed dose-dependent toxicity against neonates of the six tested lepidopteran species in terms of mortality and growth inhibition ($P < 0.05$, Tables S1 and S2). The lepidopteran pest species tested differed in their levels of susceptibility to UFT038 and the reference HD-1 *Bt* strains (Figs. 2a and 2b). However, the magnitude of toxicity was comparable for both *Bt* strains in most cases (Fig. 2a and 2b, Tables S1, and S2). Based on overlapping 95% confidence limits, UFT038 and HD-1 were equally toxic against to *S. cosmioides*, *S. frugiperda* (susceptible population), *H. zea*, *H. armigera*, and *S. frugiperda* R_Cry1+2Ab. On the other hand, non-overlapping limits support HD-1 was significantly more toxic to *S. eridania*, *C. includens*, and Cry1F-resistant *S. frugiperda* (Table S1). In contrast, growth inhibition non-overlapping confidence limits supported that UFT038 was more toxic than HD-1 to *S. cosmioides* and *S. frugiperda* LabSS, while HD-1 was more active against *C. includens* and Cry1F-resistant *S. frugiperda* (Fig. 2b, Table S2).

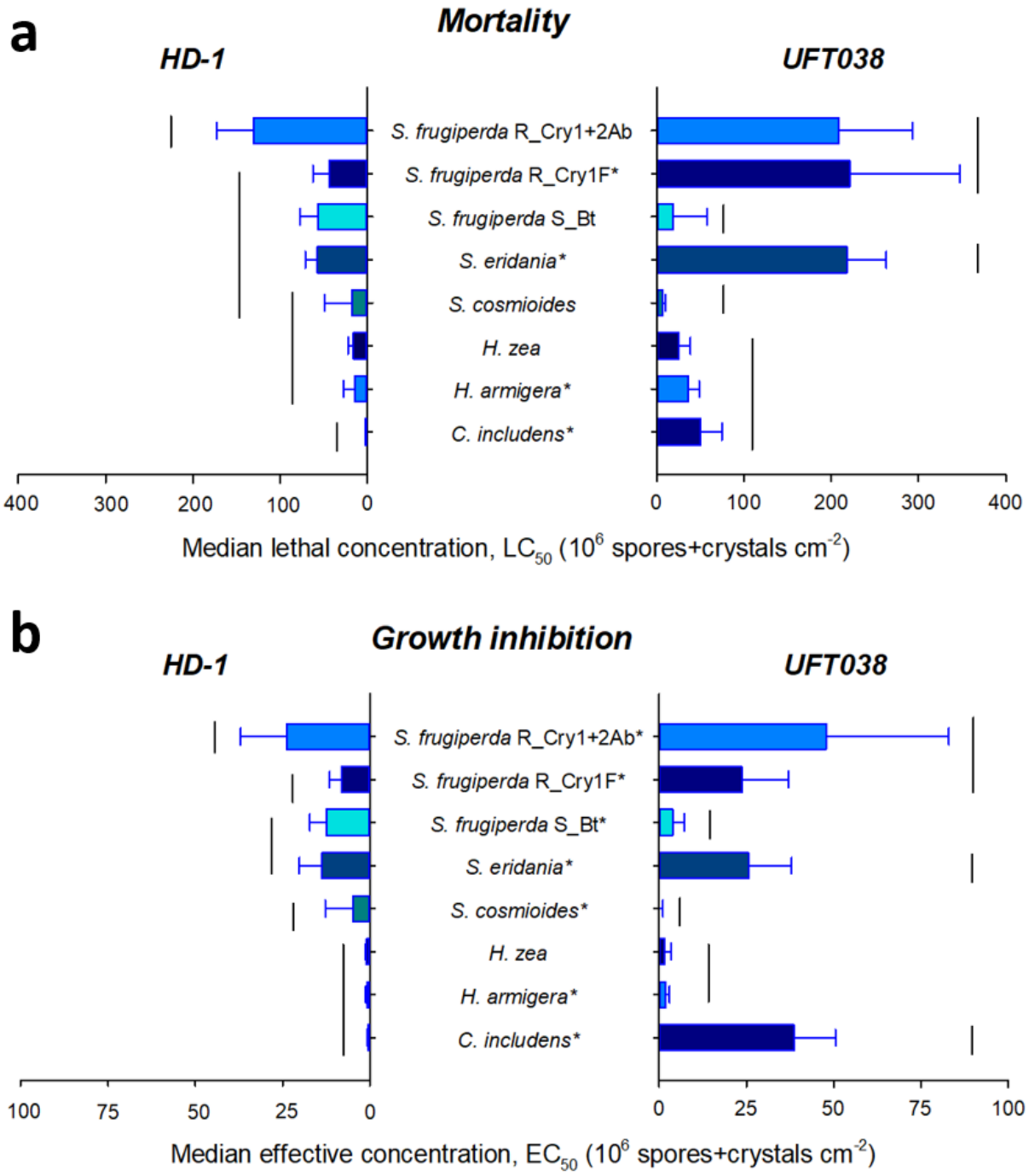

Fig. 2 Comparative potency of the UFT038 and HD-1 *Bt* strains against six lepidopteran pest species and two *Bt*-resistant populations of the fall armyworm (*Spodoptera frugiperda*). (**a**) Toxicity based on larval mortality (i.e., lethal concentration values). (**b**) Toxicity based on larval growth inhibition (i.e., effective concentration values). While a same-level-line segment indicates insect species or populations that are not significantly different ($P > 0.05$) in susceptibility to the particular *Bt* strain, asterisks indicate a significant difference ($P < 0.05$) between UFT038 and HD-1 strains against a particular insect species or population (Robertson et al. 2017). Bioassay results and statistical analysis are shared in Tables S1 and S2.

3.2 Characterization of the UFT038 genome

The draft genome of *Bt* UFT038 comprises a ~ 5,631,810 bp circular chromosome with 35.3% GC content and 5,469 predicted protein-coding regions (CDs) (Table 1). A total of 11 plasmid-like sequences of 2,062–293,573 bp were identified (Table 1).

Table 1 Draft genome features of *Bt* UFT038 strain.

Replicon	Mean coverage	Gapped sites (%)	Estimated size (bp)	GC (%)	CDS (n°)	Pesticidal proteins gene*
chromosome	1.8	9.7	5,631,810	35.3	5,469	-
p2062	82.9	-	2,062	34.8	3	-
p7635	104.8	-	7,635	32.2	9	-
p8509	103.4	0.7	8,509	31.0	9	-
p14870	51.9	1.3	14,870	40.2	28	-
p14889	73.1	0.1	14,889	31.3	23	-
p17064	78	1.0	17,064	31.5	14	-
p46634	88.5	-	46,634	35.5	68	-
p80701	44.7	5.7	80,701	33.9	86	-
p82414	47.8	3.5	82,414	33.9	90	-
p95988	82.1	17.9	95,988	32.1	97	*cry1Ac11*
p293574	51.6	11.4	293,574	34.0	265	*cry1Aa8*, *cry1Ia44*, *cry2Aa9*, *cry2Ab35* and *vip3Af5*

*Classification based on BtToxin_Scanner_Digger tool

The 4,771 CDs predicted in the UFT038 *Bt* chromosome were assigned to clusters of orthologous groups (COG) categories (Fig. 4). Most of the CDS mapped to category E (amino acid transport and metabolism), followed by categories K (transcription), M (Cell wall/Membrane/envelope biogenesis), P (Inorganic ion transport and metabolism), and C (Energy production and conversion) (Fig. 3).

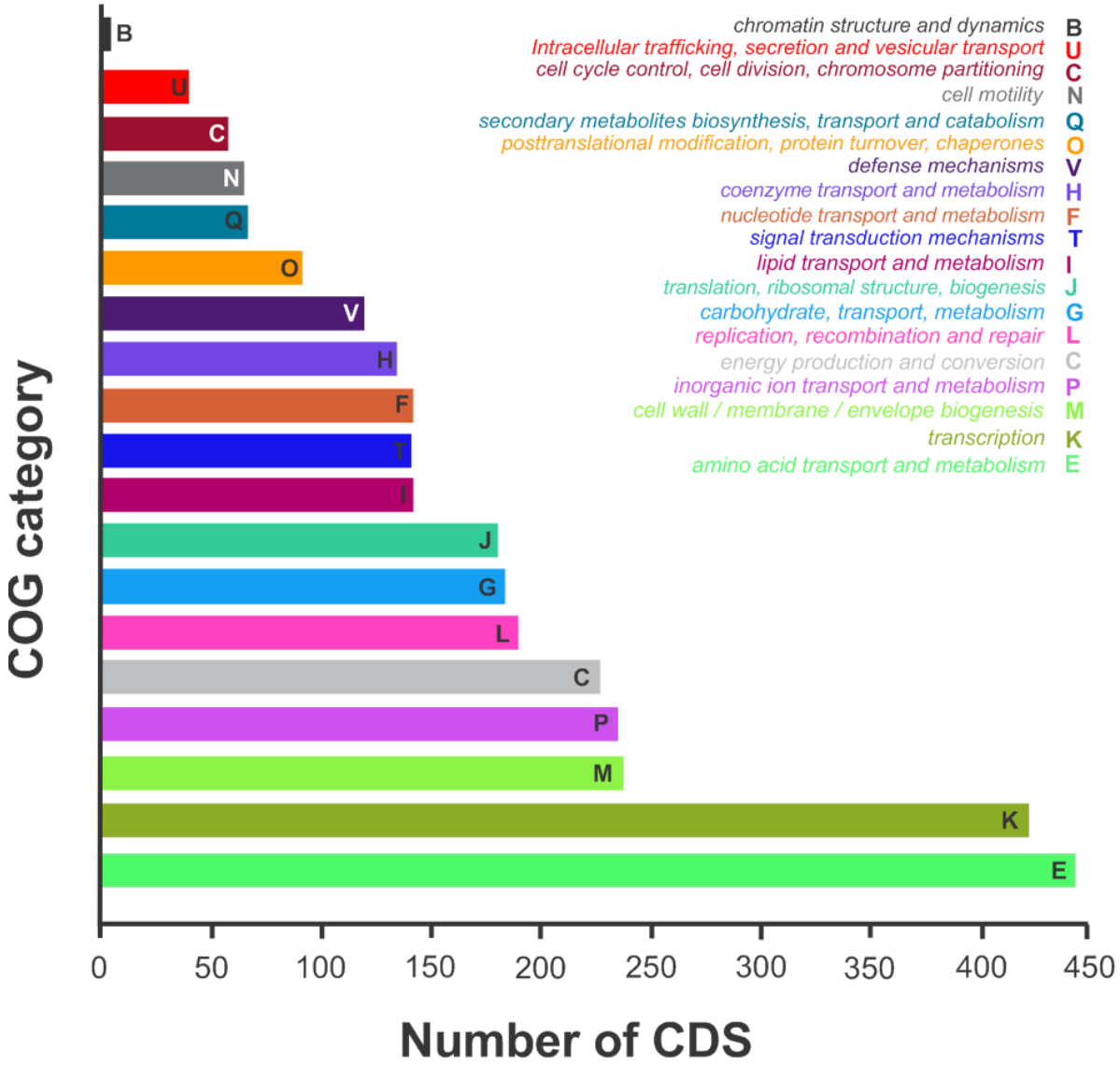

Fig. 3 CDS assigned to COG categories in the chromosome of UFT038. The x-axis shows the COG categories, and the y-axis represents the number of CDS. Queries having no COG category or belonging to category S (27.3% of all categories) were excluded from the bar plot.

Screens for virulence-factors genes in the UFT038 chromosome utilized the VFDB database as a reference. These searches detected matches to genes encoding enzymes (*InhA1*, *plcA*, *piplc*, *sph*, and *chi1-2*) and toxins (*alo*, *cytK*, *hlyIII*, *hblA-D*, and *nheA-C*) or having roles in immune evasion (polysaccharide capsule genes), iron acquisition (*dhbA-C*, *dhbE*, *dhbF*, *hal*, *ilsA*, and *asbA-F*), regulation (*pagR-XO2*, *papR*, *plcR*, and *lisR*), urease production (*ureB*), magnesium uptake (*mgtB*), and O-antigen determination (*ddhA*, *wcaG1*) (Table 2).

Table 2 Summary of virulence factors identified in UFT038 chromosome.

VFclass	Virulence factors	Related genes	UFT038 (Prediction)	Start-End	Lenght	Strand
Enzyme	Immune inhibitor A metalloproteinase	-	orf01456	1,480,114 - 1,482,480	2,367	Plus
			orf02924	2,883,054 - 2,885,441	2,388	Plus
		inhA	orf00679	775,663 - 778,062	2,400	Plus
	Phosphatidylcholine-preferring phospholipase C (PC-PLC)	*plcA*	orf00687	783,531 - 784,382	852	Plus
	Phosphatidylinositol-specific phospholipase C (PI-PLC)	*piplc*	orf04040	3,877,292 - 3,878,281	990	Minus
	Sphingomyelinase (SMase)	*sph*	orf00688	784,459 - 785,505	1,047	Plus
	Chitinases	*chi1*	orf02625	524,009 - 521,985	2,025	Minus
		chi2	orf19725	3,843,736 - 3,844,818	1,083	Plus
Immune evasion	Polysaccharide capsule	-	orf05689	5,372,127 - 5,372,957	831	Minus
			orf05691	5,373,056-5,373,724	669	Minus
			orf05692	5,374,947 - 5,376,269	1,323	Minus
			orf05693	5,376,266 - 5,377,120	855	Minus
			orf05702	5,383,431 - 5,384,324	894	Minus
			orf05703	5,384,554 - 5,385,330	777	Minus
			orf05704	5,385,433 - 5,386,134	702	Minus
			orf05705	5,386,124 - 5,386,867	744	Minus
			orf05707	5,387,130 - 5,387,903	774	Minus
Iron acquisition	Bacillibactin	*dhbA*	orf02489	2,462,079 - 2,462,864	786	Plus
		dhbB	orf02492	2,465,743 - 2,466,636	894	Plus
		dhbC	orf02490	2,462,890 - 2,464,089	1,200	Plus
		dhbE	orf02491	2,464,102 - 2,465,718	1,617	Plus
		dhbF	orf02493	2,466,670 - 2,473,827	7,158	Plus
	Hal	*hal*	orf00562	632,622 - 633,266	645	Plus
	IlsA	*ilsA*	orf00563	633,277 - 635,616	2,340	Plus
			orf01506	1,527,026 - 1,529,326	2,301	Plus
	Petrobactin	*asbA*	orf02093	2,077,116 - 2,078,924	1,809	Plus
		asbB	orf02094	2,078,997 - 2,080,823	1,827	Plus
		asbC	orf02095	2,080,810 - 2,082,048	1,239	Plus
		asbD	orf02096	2,082,072 - 2,082,311	240	Plus
		asbE	orf02097	2,082,335 - 2,083,318	984	Plus
		asbF	orf02098	2,083,356 - 2,084,198	843	Plus
Regulation	PagR-XO2	*pagR-XO2*	orf02660	2,635,113 - 2,635,406	294	Plus
	PlcR-PapR quorum sensing	*papR*	orf05785	5,468,932 - 5,471,670	2,739	Plus
			orf05786	5,471,690 - 5,471,836	147	Minus
		plcR	orf05787	5 471 925 - 5,472,782	858	Minus
	LisR	*lisR*	orf03294	3,201,438 - 3,202,112	675	Minus

Toxin	Anthrolysin O/Cereolysin O/Hemolysin I	*alo*	orf05512	5,192,645 - 5,194,183	1,539	Minus
	Cytotoxin K (Hemolysin IV)	*cytK*	orf01196	1,254,413 -1,255,423	1,011	Minus
	Hemolysin III homolog	*Undetermined*	orf05894	5,591,53 - 5,592,193	663	Minus
	Hemolysin III	*hlyIII*	orf02379	2,358,924 - 2,359,640	717	Minus
	Hemolytic enterotoxin HBL	*hblA*	orf02668	2,641,670 - 2,642,890	1,221	Plus
			orf02787	2,759,892 - 2,761,034	1,143	Plus
			orf02788	2,761,410 - 2,762,810	1,401	Plus
		hblC	orf02666	2,639,018 - 2,640,412	1,395	Plus
		hblD	orf02667	2,640,409 - 2,641,638	1,230	Plus
			orf02786	2,758,650 - 2,759,870	1,221	Plus
	Non-hemolytic enterotoxin (Nhe)	*nheA*	orf01993	1,981,381 - 1,982,541	1,161	Plus
		nheB	orf01994	1,982,579 - 1,983,787	1,209	Plus
		nheC	orf01995	1,983,886 - 1,984,974	1,089	Plus
Acid resistance	Urease	*ureB*	orf03906	3,761,327 – 3,763,027	1,701	Minus
Magnesium uptake	Mg^{2+} transport	*mgtB*	orf04453	4,271,123 – 4,273,849	2,727	Plus
Others	O-antigen	*ddhA*	orf03588	3,455,157 – 3,455,924	768	Plus
		wcaG1	orf03757	3,606,692 – 3,607,780	1,089	Plus

The p95988 megaplasmid contains the *cry1Ac11* gene, while the p293574 plasmid contains *cry1Aa8*, *cry1Ia44*, *cry2Aa9*, *cry2Ab35*, and *vip3Af5* toxin genes (Table 1, Fig. 4a and 4b).

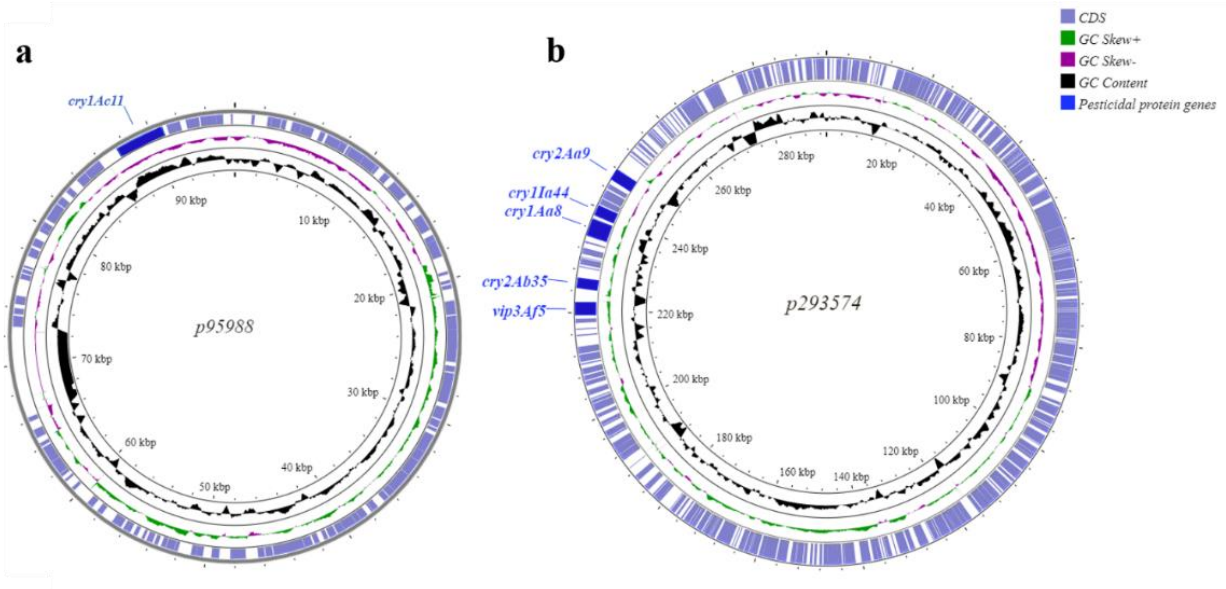

Fig. 4 Pesticidal proteins genes regions detected in megaplasmids of the UFT038. (**a**) p95988 and (**b**) p293573 megaplasmids. The genes were classified using BtToxin_Digger tool.

3.3 MLST and phylogeny

In the MLST analysis, the number of alleles per locus was: 33 (*glp*), eight (*gmk*), 13 (*ilvD*), 19 (*pta*), two (*pur*), 17 (*pycA*), and 104 (*tpi*). These alleles were assigned to different STs, which correlated with *B. cereus*, *Bt*, and *Bt* with multiple serovars (2–4 different serovars), including ST8 correlated with *Bt kurstaki*. For further evidence of serotyping, the phylogenetic analysis using the *gyrB* gene sequence showed that the UFT038 strain clustered with representative *Bt kurstaki* isolates (Fig.5).

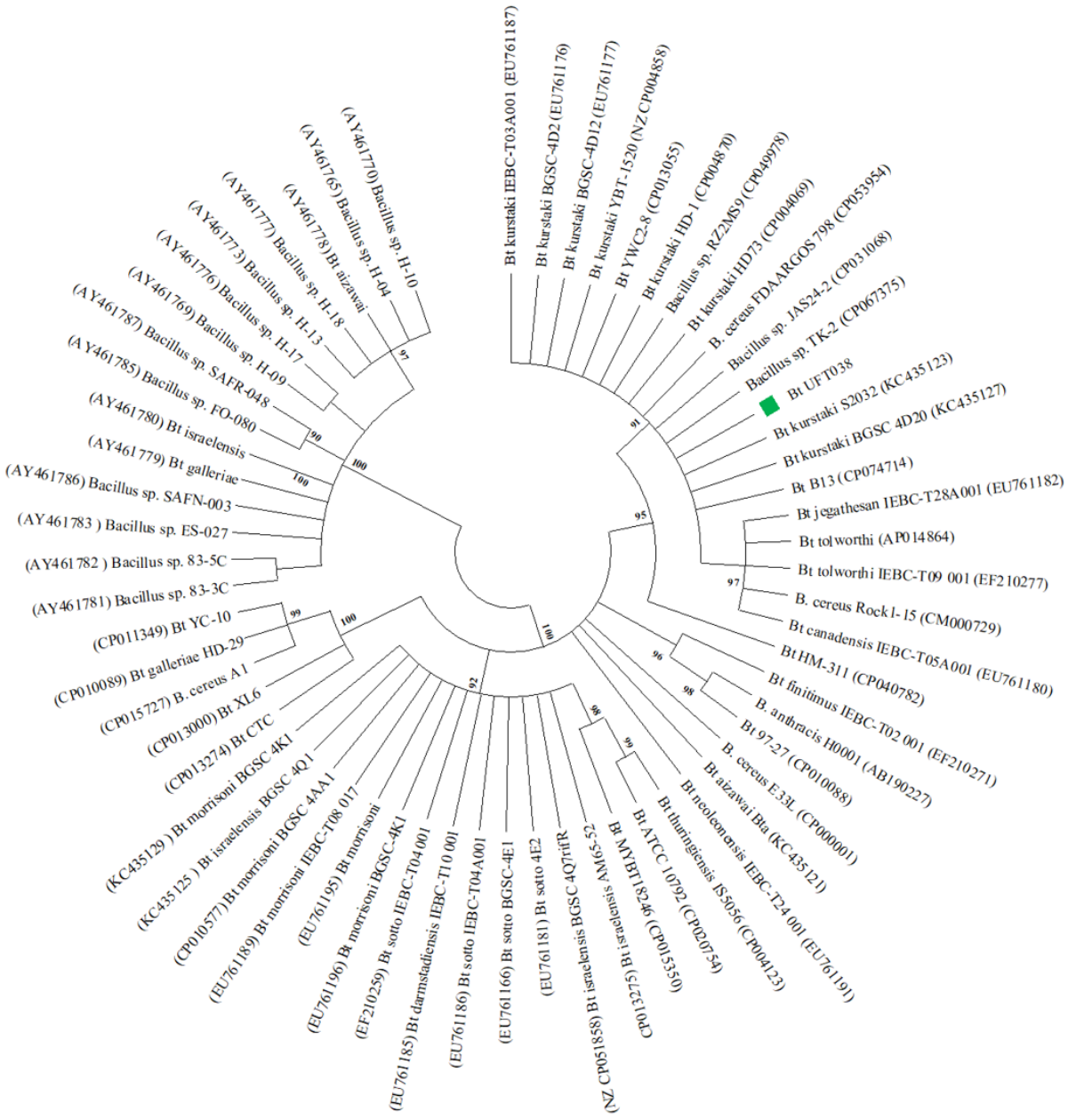

Fig.5 Phylogenetic tree constructed based on *gyrB* gene sequences of the UFT038 and other *Bacillus* ssp. strains. Bootstrap values (%) presented at the branches were calculated from 1,000 replications. Bootstrap values <90% were disregarded for branches in the graphical representation of the tree.

4 Discussion

Considering that there is an increasing number of reports describing populations of lepidopteran pests with high levels of resistance to insecticidal *Bt* proteins in transgenic crops or formulated biorational insecticides, the search for novel *Bt* isolates with different modes of action is of great relevance for properly integrated pest management. Here, we present a novel Neotropical *Bt* strain (UFT038) isolated from Northern Brazil soil samples and exhibited adequate toxicity to six lepidopteran pest species, including two fall armyworms (*S. frugiperda*) populations that are already resistant to conventional *Bt* toxins. Our phylogenetic analysis revealed that the UFT038 strain clustered with known strains of *Bt* serovar *kurstaki* and produced bipyramidal and cuboid crystals during sporulation.

Insecticidal proteins and virulence factors from UFT038 showed a relatively wide range of toxicity against various lepidopteran species, comparable to that of the HD-1 reference *Bt* strain. The broad toxicity observed is likely associated with the presence of diverse insecticidal proteins, including three Cry proteins and one Vip3A protein. The Cry1, Cry2, and Vip3A proteins represent different modes of action as they do not share receptor sites in tested species (Jakka et al., 2015). Pyramiding of *Bt* insecticidal proteins with a distinct mode of action (i.e., binding to different midgut receptors) is a recognized strategy to increase activity and delay the selection of resistant individuals (Bengyella et al., 2018). Similarly, the presence of insecticidal protein genes differing in the mode of action is also desirable in *Bt* strains amenable to field application.

The UFT038 genome included *cry* genes encoding insecticidal proteins considered effective against all insect species tested. Larvae of *Spodoptera* spp. are one of the most economically relevant pests of many crops globally. An emerging pest, *S. cosmioides*, has increasingly become a more significant pest causing damage to several crops such as cotton, soybeans, tomato, and cowpea. Compared with the reference HD-1 strain, UFT038 was more toxic to *S. cosmioides*. Santos et al. (2009) previously reported the isolation of another *Bt* strain was more harmful to the same species when compared with the reference strain. This difference in toxicity among *Bt* strains highlights the importance of screening for strains controlling emerging insect pests that are also not well controlled by specific current *Bt* transgenic cultivars (Rabelo et al., 2020a). The toxicity of UFT038 to *S. cosmioides* was higher than that of *S. eridania*, another emergent lepidopteran pest (Rabelo et al., 2020b).

The evolution of resistance threatens the sustainability of transgenic *Bt* crops for pest control (Tabashnik et al. 2017). When studied, practical resistance to *Bt* crops is associated

with mutations affecting receptor genes for the insecticidal proteins produced by the plant (Jurat-Fuentes et al., 2021). Thus, available molecular reports on resistance mechanisms to Cry1F corn in *S. frugiperda* conclude it is receptor-mediated (Boaventura et al., 2020)·. Consequently, strains of *S. frugiperda* with resistance to Cry1F corn remain susceptible to *Bt* strains producing multiple insecticidal proteins not sharing receptors with Cry1F (Jakka et al. 2014). Interestingly, the UFT038 strain displayed slightly lower activity than HD-1 against *S. frugiperda* strains resistant to Cry1F (almost 3-fold) or Cry1A.105-Cry2Ab (2-fold but not significant based on overlapping confidence intervals) while being more active (3-fold) against a reference susceptible *S. frugiperda* strain. Analysis of the UFT038 genome detected six insecticidal protein genes (*cry1Ac11*, *cry1Aa8*, *cry1Ia44*, *cry2Aa9*, *cry2Ab35*, and *vip3Af5*), while HD-1 produces a mixture of four Cry insecticidal proteins (Cry1Aa, Cry1Ab, Cry1Ac, Cry2Aa, Cry2Ab and Cry1Ia) (Caballero et al., 2020) and Vip3Aa (Milne et al., 2008).

One plausible explanation for the significant discrepancy between UFT038 and HD-1 activity on Cry1F-resistant *S. frugiperda* may be the expression of different levels of insecticidal proteins overcoming Cry1F resistance, such as Cry2A and Vip3A proteins (Vélez et al., 2013). Higher-level production of inactive proteins in Cry1F-resistant *S. frugiperda*, such as Cry1Ac and Cry1Aa, can explain UFT038 being more active than HD-1 against susceptible *S. frugiperda*. Since this study used mixtures of spores and crystals in the toxicity bioassays, some spore virulence factors may also be involved in the toxicity of the UFT038 strain (Liu et al., 1998; Dubovskiy et al., 2021).

The UFT038 chromosome harbors multiple virulence factor genes that may facilitate colonization and pathogenesis (Vilas-Bôas et al., 2012). Among the virulence factors detected in the UFT038 chromosome, those encoded by *inhA* genes may help neutralize the host immune system thought to hydrolyze antibacterial proteins in the hemolymph (Guillemet et al., 2010; Pohare et al., 2021). Other virulence genes have hemolytic and pore-formation activity (*hlyIII*, *cytK*, *hblA*, *hblC*, and *hblD*) (Vilas-Bôas et al., 2012) and cytotoxicity (*nheA*, *nheB*, *nheC*) (Swiecicka et al., 2006; Vilas-Bôas et al., 2007; Vilas-Bôas et al., 2012).

In conclusion, this study shows the potential that the UFT038 strain has to be exploited to develop novel biopesticides against lepidopteran insect pests with high toxicity to *S. cosmioides*. The genome analysis indicated that the strain carries genes related to pesticidal proteins and virulence factors.

5 References

Sree K S, Varma A (2015) An Introduction to Entomopathogenic Microorganisms. In: Sree KS, and Varma A, (eds) Biocontrol of Lepidopteran Pests: Use of Soil Microbes and their Metabolites. Springer Switzerland pp 1-10.

Cordeiro EMG, Pantoja-Gomez LM, de Paiva JB, Nascimento ARB, Omoto C, Michel AP,Correa AS (2020) Hybridization and introgression between *Helicoverpa armigera* and *H. zea*: an adaptational bridge. BMC Evol Biol 20:61. https://doi.org/10.1186/s12862-020-01621-8.

Freitas JGd, Takahashi TA, Figueiredo LL, Fernandes PM, Camargo LF, Watanabe IM, Foerster LA, Fernandez-Triana J, Shimbori EM (2019) First record of *Cotesia scotti* (Valerio and Whitfield, 2009) (Hymenoptera: Braconidae: Microgastrinae) comb. nov. parasitising *Spodoptera cosmioides* (Walk, 1858) and *Spodoptera eridania* (Stoll, 1782) (Lepidoptera: Noctuidae) in Brazil. Rev Bras Entomol 63:238-244. https://doi.org/10.1016/j.rbe.2019.05.001

Palma J, Maebe K, Guedes JVC, Smagghe G (2015) Molecular variability and genetic structure of *Chrysodeixis includens* (Lepidoptera: Noctuidae), an important soybean defoliator in Brazil. PloS ONE 10:e0121260. https://doi.org/10.1371/journal.pone.0121260

Santos KB, Neves P, Meneguim AM, dos Santos RB, dos Santos WJ, Boas GV, Dumas V, Martins E, Praça LB, Queiroz P (2009) Selection and characterization of the *Bacillus thuringiensis* strains toxic to *Spodoptera eridania* (Cramer), *Spodoptera cosmioides* (Walker) and *Spodoptera frugiperda* (Smith) (Lepidoptera: Noctuidae). Biol Control 50:157-163. https://doi.org/10.1016/j.biocontrol.2009.03.014

Specht A, Sosa-Gómez DR, Paula-Moraes SVd, Yano SAC (2013) Morphological and molecular identification of *Helicoverpa armigera* (Lepidoptera: Noctuidae) and expansion of its occurrence record in Brazil. Pesq Agrop Bras 48:689-692. https://doi.org/10.1590/S0100-204X2013000600015

Marrone PG (2019) Pesticidal natural products - status and future potential. Pest Manag Sci 75:2325-2340.

Deng C, Peng Q, Song F, Lereclus D (2014) Regulation of *cry* gene expression in *Bacillus thuringiensis*. Toxins 6:2194-2209. https://doi.org/10.3390/toxins6072194

Méric G, Mageiros L, Pascoe B, Woodcock DJ, Mourkas E, Lamble S, Bowden R, Jolley KA, Raymond B, Sheppard SK (2018) Lineage-specific plasmid acquisition and the evolution of specialized pathogens in *Bacillus thuringiensis* and the *Bacillus cereus* group. Mol Ecol 27:1524-1540. https://doi.org/10.1111/mec.14546

Zheng J, Gao Q, Liu L, Liu H, Wang Y, Peng D, Ruan L, Raymond B, Sun M, Parkhill J (2017) Comparative genomics of *Bacillus thuringiensis* reveals a path to specialized exploitation of multiple invertebrate hosts. mBio 8:e00822-00817. https://doi.org/10.1128/mBio.00822-17

Romeis J et al (2019) Genetically engineered crops help support conservation biological control. Bio Control 130: 136-154.

Tabashnik BE, Carrière Y (2017) Surge in insect resistance to transgenic crops and prospects for sustainability. Nat Biotechnol 35:926-935. https://doi.org/10.1038/nbt.3974

Boonmee K, Thammasittirong SN-R, Thammasittirong A (2019) Molecular characterization of lepidopteran-specific toxin genes in *Bacillus thuringiensis* strains from Thailand. 3 Biotech 9:117. https://doi.org/10.1007/s13205-019-1646-3

Djenane Z, Lázaro-Berenguer M, Nateche F, Ferré J (2020) Evaluation of the toxicity of supernatant cultures and spore–crystal mixtures of *Bacillus thuringiensis* strains isolated from Algeria. Curr Microbiol 77:2904-2914. https://doi.org/10.1007/s00284-020-02110-3

Baragamaarachchi RY et al (2019) Identification of a native *Bacillus thuringiensis* strain from Sri Lanka active against Dipel-resistant *Plutella xylostella*. PeerJ 7: e7535.

Gomis-Cebolla J, Scaramal Ricietto AP, Ferré J (2018) A genomic and proteomic approach to identify and quantify the expressed *Bacillus thuringiensis* proteins in the supernatant and parasporal crystal. Toxins 10:193. https://doi.org/10.3390/toxins10050193

Zhu L, Peng D, Wang Y, Ye W, Zheng J, Zhao C, Han D, Geng C, Ruan L, He J, Yu Z, Sun M (2015) Genomic and transcriptomic insights into the efficient entomopathogenicity of *Bacillus thuringiensis*. Sci Rep 5:14129. https://doi.org/10.1038/srep14129

Stewart G, Johnstone K, Hagelberg E, Ellar DJ (1981) Commitment of bacterial spores to germinate a measure of the trigger reaction. Biochem J 198:101-106. https://doi.org/10.1042/bj1980101

Azzouz H, Kebaili-Ghribi J, ben Farhat-Touzri D, Daoud F, Fakhfakh I, Tounsi S, Jaoua S (2014) Selection and characterisation of an HD1-like *Bacillus thuringiensis* isolate with a high insecticidal activity against *Spodoptera littoralis* (Lepidoptera: Noctuidae). Pest Manag Sci 70:1192-1201. https://doi.org/10.1002/ps.3661

Zouari N, Achour O, Jaoua S (2002) Production of delta-endotoxin by *Bacillus thuringiensis* subsp *kurstaki* and overcoming of catabolite repression by using highly concentrated gruel and fish meal media in 2-and 20-dm3 fermenters. J Chem Technol Biotechnol 77:877-882. https://doi.org/10.1002/jctb.650

Santos-Amaya OF, Tavares CS, Rodrigues JVC, Campos SO, Guedes RNC, Alves AP, Pereira EJG (2017) Fitness costs and stability of Cry1Fa resistance in Brazilian populations of *Spodoptera frugiperda*. Pest Manag Sci, 73:35-43. https://doi.org/10.1002/ps.4312 (a)

Santos-Amaya OF, Tavares CS, Rodrigues JVC, Souza TC, Rodrigues-Silva N, Guedes RNC, Alves AP, Pereira EJG (2017) Magnitude and allele frequency of Cry1F resistance in field populations of the fall armyworm (Lepidoptera: Noctuidae) in Brazil. J Econ Entomol 110:1770-1778. https://doi.org/10.1093/jee/tox146 (b)

Santos-Amaya OF, Rodrigues JVC, Souza TC, Tavares CS, Campos SO, Guedes RNC, Pereira EJG (2015) Resistance to dual-gene *Bt* maize in *Spodoptera frugiperda*: Selection, inheritance, and cross-resistance to other transgenic events. Sci Rep 5:18243-18243. https://doi.org/10.1038/srep18243

Rabelo MM, Matos JML, Orozco-Restrepo SM, Paula-Moraes SV, Pereira EJG (2020) Like parents, like offspring? Susceptibility to *Bt* toxins, development on dual-gene *Bt* cotton, and parental effect of Cry1Ac on a non-target lepidopteran pest. J Econ Entomol 113:1234-1242. 10.1093/jee/toaa051 (a)

Rabelo MM, Matos JML, Santos-Amaya OF, França JC, Gonçalves J, Paula-Moraes SV, Guedes RNC, Pereira EJG (2020) *Bt*-toxin susceptibility and hormesis-like response in the invasive southern armyworm (*Spodoptera eridania*). J Crop Prot 132:105129-105129. 10.1016/j.cropro.2020.105129 (b)

Kasten P, AACM P, Parra JRP (1978) Dados Biologicos Comparativos de *Spodoptera frugiperda* (J.E. Smith, 1977) em duas dietas artificiais e substrate natural. Revista Agricola 53:68-78.

Greene GL, Leppla NC, Dickerson WA (1976) Velvetbean caterpillar: a rearing procedure and artificial medium. J Econ Entomol 69:487-488. https://doi.org/10.1093/jee/69.4.487

Finney DJ (1971) Statistical methods in biological assay. 2 nd ed. London: Griffin 333p.

Robertson JL, Jones MM, Olguin E, Alberts B (2007) Bioassays with Arthropods: CRC press.

Andrews S (2015) FastQC—A Quality Control Tool for High Throughput Sequence Data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed 09 April 2022.

Bushnell B (2014) BBMap: a fast, accurate, splice-aware aligner. Join Genome Institute, Departament of Energy: Lawrence Berkeley National Lab. (LBNL), Berkeley, CA (United States).

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647-1649. https://doi.org/10.1093/bioinformatics/bts199

Gaafar YZA, Ziebell H (2020) Comparative study on three viral enrichment approaches based on RNA extraction for plant virus/viroid detection using high-throughput sequencing. PloS ONE 15:e0237951. https://doi.org/10.1371/journal.pone.0237951

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455-477. https://doi.org/10.1089/cmb.2012.0021

Galata V, Fehlmann T, Backes C, Keller A (2019) PLSDB: a resource of complete bacterial plasmids. Nucleic Acids Res 47:D195-D202. https://doi.org/10.1093/nar/gky1050

Wuyts S, Allonsius CN, Wittouck S, Thys S, Lievens B, Weckx S, De Vuyst L, Sarah L (2019) Comparative genome analysis of *Lactobacillus mudanjiangensis*, an understudied member of the *Lactobacillus plantarum* group. Microb Genom 5. https://doi.org/10.1099/mgen.0.000286

Liu H, Zheng J, Bo D, Yu Y, Ye W, Peng D, Sun M (2020) BtToxin_Digger: a comprehensive and high-throughput pipeline for mining toxin protein genes from *Bacillus thuringiensis*. BioRxiv. https://doi.org/10.1101/2020.05.26.114520

Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J (2021) eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. BioRxiv 38: 5825-5829 https://doi.org/10.1093/molbev/msab293

Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ (2019) eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res 47:D309-D314. https://doi.org/10.1093/nar/gky1085

Liu B, Zheng D, Jin Q, Chen L, Yang J (2019) VFDB 2019: a comparative pathogenomic platform with an interactive web interface. Nucleic Acids Res 47:D687-D692. https://doi.org/10.1093/nar/gky1080

Tamura, K, Stecher, G, Kumar, S (2011) MEGA11: molecular evolutionary genetics analysis version 11. Molecular biology and evolution 38:3022-3027. https://doi.org/10.1093/molbev/msab120

Jakka S et al (2015) Cry toxin binding site models and their use in strategies to delay resistance evolution." *Bt* Resistance: Characterization and Strategies for GM Crops Producing *Bacillus thuringiensis* Toxins, eds Soberón M, Gao Y, Bravo A (Centre for Agriculture and Biosciences International, Oxfordshire, UK) 138-149.

Bengyella L et al (2018) Global challenges faced by engineered *Bacillus thuringiensis* Cry genes in soybean (Glycine max L.) in the twenty-first century. 3 Biotech 8: 464.

Jurat-Fuentes JL et al (2021) Mechanisms of resistance to insecticidal proteins from *Bacillus thuringiensis*. Annu Rev Entomol 66: 121-140.

Boaventura D et al (2020) Molecular characterization of Cry1F resistance in fall armyworm, *Spodoptera frugiperda* from Brazil." Insect Biochem Mol Biol 116: 103280.

Jakka SR et al (2014) *Spodoptera frugiperda* (J. E. Smith) with field-evolved resistance to *Bt* maize are susceptible to *Bt* pesticides. J. Invertebr. Pathol 122: 52-54.

Caballero J et al (2020) Unraveling the composition of insecticidal crystal proteins in *Bacillus thuringiensis*: a proteomics approach. Appl. Environ. Microbiol 86: e00476-20. 10.1128/AEM.00476-20

Milne R, Liu Y, Gauthier D, Van Frankenhuyzen K (2008) Purification of Vip3Aa from *Bacillus thuringiensis* HD-1 and its contribution to toxicity of HD-1 to spruce budworm (*Choristoneura fumiferana*) and gypsy moth (Lymantria dispar) (Lepidoptera). J Invertebr Pathol99:166-172. https://doi.org/10.1016/j.jip.2008.05.002

Vélez AM et al (2013) Inheritance of Cry1F resistance, cross-resistance and frequency of resistant alleles in *Spodoptera frugiperda* (Lepidoptera: Noctuidae). Bull Entomol Res 103: 700-713.

Liu YB et al (1998) Synergism between *Bacillus thuringiensis* spores and toxins against resistant and susceptible diamondback moths (*Plutella xylostella*). Appl Environ Microbiol 64:1385-1389.

Dubovskiy IM et al (2021) *Bacillus thuringiensis* spores and Cry3A toxins act synergistically to expedite Colorado potato beetle mortality." Toxins (Basel) 13(11).

Vilas-Bôas GT, Alvarez RC, dos Santos CA, Vilas-Boas LA (2012) Fatores de virulência de *Bacillus thuringiensis*: o que existe além das proteínas Cry. EntomoBrasilis 5:1-10. https://doi.org/10.12741/ebrasilis.v5i1.146

Guillemet E, Cadot C, Tran S-L, Guinebretière M-H, Lereclus D, Ramarao N (2010) The InhA metalloproteases of *Bacillus cereus* contribute concomitantly to virulence. J Bacteriol 192:286-294. https://doi.org/10.1128/JB.00264-09

Pohare MB, Wagh SG, Udayasuriyan V (2021) *Bacillus thuringiensis* as Potential Biocontrol Agent for Sustainable Agriculture. In: Yadav AN, Singh J, Singh C, and Yadav N, (eds) Current Trends in Microbial Biotechnology for Sustainable Agriculture. Environmental and Microbial Biotechnology 439-468. https://doi.org/10.1007/978-981-15-6949-4_18

Swiecicka I, Van der Auwera GA, Mahillon J (2006) Hemolytic and nonhemolytic enterotoxin genes are broadly distributed among *Bacillus thuringiensis* isolated from wild mammals. Microb Ecol 52:544-551. https://doi.org/10.1007/s00248-006-9122-0

Vilas-Bôas GTV-BT, Peruca APSPPS, Arantes OMNAMN (2007) Biology and taxonomy of *Bacillus cereus*, *Bacillus anthracis*, and *Bacillus thuringiensis*. Can J Microbiol53:673-687. https://doi.org/10.1139/W07-029

Supplementary Information Table S1 Toxicity (larval mortality) of *Bt* strains (UFT038 and HD-1) against six lepidopteran species and two Cry1-Bt-resistant populations of *Spodoptera frugiperda*. Table S2 Toxicity (growth inhibition) of *Bt* strains against six lepidopteran species and two Cry1-Bt-resistant populations of *Spodoptera frugiperda*.

CHAPTER IV: Genomic-proteomic analysis of a novel *Bacillus thuringiensis* strain: toxicity against two Lepidoptera pests, abundance of Cry1Ac5 toxin, and presence of InhA1 virulence factor

Published in: Archives of Microbiology. doi.org/10.1007/s00203-023-03479-y (Annex D).

Abstract

Bacillus thuringiensis (*Bt*) is a biological alternative to the indiscriminate use of chemical insecticides in agriculture. Due to resistance development on insect pests to *Bt* crops, isolating novel *Bt* strains is a strategy for screening new pesticidal proteins or strains containing toxin profile variety that can delay resistance. Besides, the combined genomic and proteomic approaches allow identifying pesticidal proteins and virulence factors accurately. Here, the genome of a novel *Bt* strain (*Bt* TOL651) was sequenced, and the proteins from the spore-crystal mixture were identified by proteomic analysis. Toxicity bioassays with the spore-crystal mixture against larvae of *Diatraea saccharalis* and *Anticarsia gemmatalis*, key pests of sugarcane and soybean, respectively, were performed. The toxicity of *Bt* TOL651 varies with the insect; *A. gemmatalis* (LC_{50} =1.45 ng.cm^{-2}) is more susceptible than *D. saccharalis* (LC_{50} = 73.77 ng.cm^{-2}). Phylogenetic analysis of the *gyrB* gene indicates that TOL651 is related to *Bt kenyae* strains. The genomic analysis revealed the presence of *cry1Aa18*, *cry1Ac5*, *cry1Ia44*, and *cry2Aa9* pesticidal genes. Virulence factors genes such as phospholipases (*plcA*, *piplc*), metalloproteases (*inhA*), hemolysins (*cytK*, *hlyIII*, *hblA*, *hblC*, *hblD*), and enterotoxins (*nheA*, *nheB*, *nheC*) were also identified. The combined use of the genomic and proteomic data indicated the expression of Cry1Aa18, Cry1Ac5, and Cry2Aa9 proteins, with the Cry1Ac5 being the most abundant. InhA1 also was expressed and may contribute to *Bt* TOL651 pathogenicity. These results provide *Bt* TOL651 as a new tool for the biocontrol of lepidopteran pests.

Keywords: Biopesticide; *Bacillus thuringiensis*; Cry proteins; Genome sequence; Proteomic; Toxicity; Lepidoptera pest

1 Introduction

Sugarcane (*Saccharum officinarum*), corn (*Zea mays*), and soybean (*Glycine max* L) are high-value crops and applied to different purposes such as food and biofuel production (de Matos et al., 2020; Heinrichs et al., 2017). The sugarcane borer, *Diatraea saccharalis* (Fabricius, 1794) (Lepidoptera: Crambidae), is a key pest of sugarcane and corn. The velvetbean caterpillar, *Anticarsia gemmatalis* (Hübner, 1818) (Lepidoptera: Noctuidae), is a key pest of the soybean (Horikoshi et al., 2022; Mendonça et al., 2020). These insects are among Brazil's most relevant Lepidoptera pests causing damage to the raw material used in food and biofuel processing (Silva, 1995; Praça et al., 2006; Moscardi et al., 2012; Dinardo-Miranda et al., 2013).

Biocontrol is a safe alternative to reduce the use of chemical insecticides in crop pest management. Entomopathogenic microorganisms have been considered important agents for this purpose with relevant importance for the bacteria *Bacillus thuringiensis* (*Bt*) (Bel et al., 2017; Daquila et al., 2019; Sanahuja et al., 2011). Biological control products based on *Bt* are being used as biopesticides for decades and currently have the majority of the biological control market share in the world (Arthur and Dara, 2019; Lacey et al., 2015; Sena da Silva et al., 2022). *Bt* is a Gram-positive, spore-forming bacterium that can produce parasporal crystal proteins during the sporulation phase (Cry and Cyt proteins) (Pohare et al., 2021). These proteins when ingested by a susceptible insect are solubilized in the alkaline midgut environment, acquiring an oligomeric form, binding to the midgut cells membrane, leading to the destruction of these cells and their insecticidal proprieties (Frankenhuyzen, 2009; Bravo et al., 2007, 2011; Pinheiro and Valicente, 2021). *Bt* can produce pesticidal proteins during the vegetative phase (Vip proteins) (Pohare et al., 2021). Virulence factors such as metalloproteases, chitinases, hemolysins, and enterotoxins also represent *Bt* pathogenicity (Malovichko et al., 2019; Palma et al., 2014).

Bt is an important biopesticide used against lepidopteran pests as spray formulations and *Bt* crops (transgenic plants that express Cry and/or Vip3 proteins) (Castro et al., 2019; Daquila et al., 2019; Horikoshi et al., 2022; Srikanth et al., 2011). However, the evolution of resistance to *Bt* crops in lepidopteran pests has been reported, including in *D. saccharalis* (de Oliveira et al.,2022; Huang et al., 2015) and *A. gemmatalis* (Pezenti et al., 2021). Therefore, the isolation of novel *Bt* strains is an important strategy for the discovery of new pesticidal proteins or strains containing a range of toxin profiles that can delay the target insect's resistance.

Characterizing novel *Bt* strains and studying the genome and proteome is important to understand their pathogenicity. Genome sequencing technology has accelerated the discovery of novel pesticidal proteins, secondary metabolites, and virulence factors in *Bt* (Zghal et al. 2018; Liu et al. 2017; Cardoso et al., 2020; Day et al., 2014; Jeong et al., 2017; Jia et al., 2016; Liu et al., 2015). However, considering not all of the coding regions predicted from the annotated genome sequence are expressed, and there are cryptic pesticidal proteins (Quan et al., 2016; Rang et al., 2015). So, the pesticidal proteins expression profile could be explored using proteomic analysis. Thus, in combination with genomic studies, proteomic analysis allows the accurate identification of pesticidal proteins and virulence factors in different *Bt* strains (Baragamaarachchi et al., 2019; Gomis-Cebolla et al., 2018; Khorramnejad et al., 2020; Wu et al., 2011). Furthermore, from the genomic-proteomic analysis is also possible to estimate the protein abundance in purified parasporal crystals and spores-crystals mixtures (Baragamaarachchi et al., 2019; Huang et al., 2012; Khorramnejad et al., 2020).

This study sequenced the genome of a novel strain *Bt* TOL651, toxic against *A. gemmatalis* and *D. saccharalis*, and the entomopathogenic characteristics were explored. Additionally, a LC-MS/MS analysis of the spores-crystals mixture was performed to determine the expressed proteins.

2 Materials and Methods

2.1 Culturing of TOL651 strain

Bacillus-like colonies were isolated from soil samples collected in Tocantins state (Brazil) (11º43’45’’ S; 49º04'07" W) according to the previously described protocol (Monnerat et al., 2001). To screen for *Bt* strains, Petri dishes with a selective NYSM medium (Nutrient Yeast Extract Salt Medium) [8 g/l of nutrient broth (Difco, USA), 0.103 g/l of $CaCl_2 \cdot 2H_2O$, 0.01 g/l of $MnCl_2 \cdot 4H2O$, 0.203 g/l of $MgCl2 \cdot 6H_2O$] (Kalfon et al., 1983) containing 100 mg/L penicillin G were used and grown for 24 h at 30 ± 0.5°C at 180 rpm. Then, each colony was individually analyzed and identified by phase-contrast microscopy (× 1000) to verify the presence of inclusion bodies and crystals (Frankland and Frankland, 1887)**.** The *Bt* TOL651 was selected among 87 crystal-forming *Bt* strains (87 strains from 2.445 *Bacillus*-like colonies), due to being the most toxic among different isolated *Bt* strains and tested simultaneously against *Diatrea saccharalis* and *Anticarsia gemmatalis* in the selective bioassays, according to Monnerat et al. (2007). *Bt* HD-1was isolated from the commercial sample (Dipel WP 32 g/kg, Sumitomo Chemical do Brasil Representações Ltda.,

SP) and used as a reference strain (Cerqueira et al., 2017; Lazart et al., 2021; Sathyan et al., 2022). *Bt* HD-1 has been designated as the primary U.S. reference standard strain for toxicity against lepidopteran insects (Dulmage, 1973). *Bt* TOL651 strain was cultured at 28 °C for 12h on solid Luria-Bertani medium (LB) (10 gL^{-1} tryptone, 5 gL^{-1} yeast extract, 10 gL^{-1} NaCl, and 20 gL^{-1} Agar). A single bacterial colony was inoculated in the LB liquid medium at 28 °C, 200 rpm for 16 h, used as a starter culture for spore-crystal mixture production and in the genomic DNA extraction step.

2.2 SDS-PAGE analysis of Cry proteins

For sporulation and crystal production, a starter culture was transferred to CCY medium (30 ml) (13 mM KH_2P04, 26 mM K_2HP04, 0.002% [w/v] L-glutamine, 0.1% [w/v] casein hydrolysate, 0.1 % [w/v] bacto casitone, 0.04% bacto yeast extract, 0.6% [w/v] glycerol, 0.05 M $ZnCl_2$, 0.5 M $MgCl_2$, 0.01 M $MnCI_2$, 0.2 M $CaCl_2$, 0.05 M $FeCl_3$) and incubate at 28 °C, 200 rpm for 72h. For SDS-PAGE analysis, the crystals were purified using hexane and low-speed centrifugation, according to Mounsef et al. (2014). The spores and crystals were collected for centrifugation at 6000 rpm, 4°C for 10 min, and the pellet was washed twice by suspending it in saline solution (1M NaCl containing 0.01% Triton X-100) by centrifugation (6000 rpm, 4°C for 10 min). Following, the pellet was suspended in a 50 ml centrifuge tube with 27 ml of saline solution and sonicated at 100W of potency for 10 min. Then, 3 ml of hexane was added to the suspension following the centrifugation at 6000 rpm, 4°C for 10 min. This procedure was repeated three times. The pellet was washed twice with cold distilled water by centrifuge. Then, crystals were solubilized using 50 mM NaOH buffer at 30°C and quantified using Bradford reagent (Bio-Rad protein assay). Following, 7 µg of solubilized crystals were analyzed by 12% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and stained with Coomassie Brilliant Blue (R250, 0.4%).

2.3 Microscopy

The spore-crystal mixture of TOL651 was collected and diluted in sterile water. Then, 100 µL of this dilution was deposited over metallic supports to be dried for 24 h at 37°C, covered with gold for 180 s using an Emitech apparatus (model K550), and observed in a Zeiss scanning electron microscope (model DSM 962) at 10 or 20 Kv.

2.4 Toxicity against *Diatraea saccharalis* and *Anticarsia gemmatalis*

Eggs from *D. saccharalis* and *A. gemmatalis* were obtained from Biocontrole (biocontrole.com.br) and Embrapa Recursos Genéticos e Biotecnologia - CERNAGEN (Brasília, DF, Brazil), respectively. The insect's eggs were maintained under ideal rearing conditions (i.e., 26 °C ± 1 °C and 70 ± 10% RH and 10:14h (L: D) photoperiod) (Schmidt et al., 2001), on specific artificial diets prepared for *D. saccharalis* (Hensley and Hammond, 1968) and *A. gemmatalis* (Greene et al., 1976)

Bioassays against *D. saccharalis* were performed using 24 well cell culture plate (TPP, Techno Plastic Products AG), which poured 1.5 ml of diet into each well and, after solidification, aliquots 35 µL of ten spore-crystals dilutions (from 0.1 to 1000 ng/cm^2) were spread on the diet surface. Subsequently, a one-day-old second instar larvae were placed into each plate. Trays were closed with acrylic lids, keeping them under controlled conditions (26 °C ± 1 °C and 70 ± 10% RH and 10:14 [L: D] h). After 48 h, the surviving larvae were individually transferred to six-well cell culture plates containing a rearing diet. Larvae mortality was evaluated again on day 7 (Praça et al., 2004). Larvae were considered dead when stimulated and no movement was detected.

In the *A. gemmatalis* bioassays a total volume of 3 ml artificial diet was poured into 30 ml plastic cups. After solidification, aliquots (150 µl) of ten spore-crystal concentrations (from 0.1 to 1000 ng/cm^2) were applied on the diet surface and dried at room temperature. Then, ten one-day-old second instar larvae of *A. gemmatalis* were added to each cup. The cups were covered with plastic lids and kept under the same conditions described above. After 48 h the surviving larvae were placed in 30 ml cups containing a rearing diet and the mortality was assessed. Larval mortality was evaluated again on day 5 as described (da Silva et al., 2004).

All the bioassays were performed in triplicates. The commercial strain HD-1 was used as a reference and sterile water pH 7.0 was added as the control group. The lethal concentrations (LC_{50} and LC_{95}) were determined by Probit analysis (Finney, 1971) using the PoloPlus 1.0 (LeOra Software Berkeley, CA, USA).

2.5 Genome sequencing, data assembly, and annotation

Total DNA from *Bt* TOL651 strain was extracted and purified by the Wizard® Genomic DNA Purification Kit (Promega, Madison, WI, USA) following the manufacturer's instructions. DNA concentration and purity were checked by the NanoDrop™ 8000 apparatus

(ThermoFisher Scientific, Waltham, MA, USA), and then stored at -20°C until further use. Genome sequencing was performed using Illumina MiSeq technologies (paired-end application, 2x ~80 bp, average insert size of 200 bp), and the coverage of 137 X. FastQC v.0.11.9 (Andrews, 2015) was used for the reads libraries quality analyses, with reads being trimmed using Geneious v.10.2.6 (Kearse et al., 2012) (Workflow Trim and Filter, Error Probability: 0.05). The *de novo* assembly was performed using SPAdes v.3.10.0, using default parameters (Bankevich et al., 2012). The assembled contigs (=>500 bp) were run through MeDuSa v.1.6 (Bosi et al., 2015) for scaffolding, using complete genomes *Bt* YBT-1520 and *Bt* HD-1 (NCBI RefSeq NZ_CP004858 and CP004870 respectively) strains as a reference dataset. The quality of the final assembly was assessed using Quast v.5.0.2 (Gurevich et al., 2013) in standard mode, and completeness percentage, N50, and L50 values were obtained. The annotation and CDS prediction of scaffolds were performed using RASTtk 2.0 (Brettin et al., 2015).

2.6 Average Nucleotide Identity (ANI) and phylogenetic relationships

The genome similarity was assessed through ANI using JSpeciesWS (Richter et al., 2015); on Tetra Correlation Search (TCS) function for selecting related genomes. The Heatmap dendrogram was created using the Morpheus tool (https://software.broadinstitute.org/morpheus). The phylogenetic relationship analysis was performed using the *gyrB* gene (DNA gyrase subunit B) including other genes of the closely related *Bacillus* spp. strains retrieved from GenBank. The sequences were aligned using Clustal W, and a phylogenetic tree was created using MEGA 11 (Kumar et al., 2018), using the Neighbor-joining method and bootstrap percentages based on 1000 replications.

2.7 Gene identification of pesticidal proteins

Putative pesticidal proteins were determined using Blastx, through the Btoxin_Digger tool (scaffolds as a query) (Liu et al., 2021) and a customized database (CDS predicted as a query). The customized database was created through Geneious, using Add/Remove Database tool, from the *Bt* pesticidal proteins list available at the *Bt* nomenclaturewebsite(http://www.lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/toxins2.html). CDS with homology to the *Bt* pesticidal proteins were filtered using E-value 0.001 and word size 6 parameters. To check the presence of novel putative Cry sequences in TOL651 the

sequences obtained from the non-redundant protein database NCBI were used (Lazart et al., 2021).

2.8 Identification of virulence factors and antibiotic resistance genes

The potential virulence factors were predicted using the bacterial virulence factor database (VFDB) (Liu et al., 2019). The TOL651 genome was screened for antibiotic resistance factors, using the Resistance Gene Identifier (RGI), within Comprehensive Antibiotic Resistance Database (CARD) (Alcock et al., 2023), according to the parameters: Perfect, Strict, complete genes only and 95% identity nudge.

2.9 Proteomic analysis

The proteins in the spore-crystal mixture of the *Bt* TOL651 strain were identified by LC-MS/MS at the Biotech Company Veritas /Life Sciences at the University of São Paulo (USP, Ribeirão Preto, SP, Brazil). The sample of the spore-crystal mixture was washed three times in 1X PBS (Phosphate-Buffered Saline) and resuspended in 750 µL of solubilization buffer (8M urea, 0.5% Octyl-glucopyranoside (OG) and 0.05M Tris-HCL, pH 8.8). Then, the sample was sonicated (three cycles of 60 seconds, 30% amplitude, and shut off for two seconds) and maintained on ice. The solubilized proteins were quantified through the Bradford method using Protein Assay Dye Reagent Concentrate (Bio-Rad Laboratories), according to the manufacturer's instructions.

The preparation of the sample for advanced mass spectrometry consisted of three main steps: i) reduction and alkylation of proteins, ii) enzymatic digestion using trypsin, and iii) cleanup/desalting. Briefly, 50 µg of the sample was subjected to disulfide bridge reduction using 50 µg of DTT (Dithiothreitol) and 60 min of incubation at 37 °C. Then, the process was followed by alkylation using 250 µg of I.A (iodoacetamide) and 60 min at room temperature in the dark. Finally, the sample was diluted five times in Tris hydrochloride (0.05M Tris-HCl, pH 8.8) and incubated with 2 µg of trypsin (Promega, V511A) at 37 °C overnight.

Previously to the mass spectrometry application, the cleanup/desalting of the sample was performed using C18 resin (Supleco). The column was calibrated using 2% acetonitrile containing 0.1% formic acid, and the elution was performed with 50% acetonitrile. The sample was then dried in a speed vac and applied in the mass spectrometer (Termo Fisher Orbitrap Eclipse), coupled to a nanoflow Nano LC-MS/MS chromatography system (Dionex Ultimate 3000 RLSCnano System, Thermofisher). Peptides were separated in nanoEase MZ

peptide BEH C18 column (130A, 1.7 µm, 75 µm x 250 mm, Waters) 300 nL/min using a 4-50% acetonitrile gradient for 90 min. The data were obtained on MS1 in the range of M/Z 375-1500 (120,000 resolution, AGC target 1E6, maximum time injection of 100 ms). The most abundant ions were submitted to MS/MS (30% collision energy, 1.2 m/z, AGC target 1E5, 15000 resolution).

The raw data were converted to mzXML format and processed using PatternLabV (Santos et al., 2022). The data was analyzed against the database created using CDS translated of the TOL651 genome (Generate Search DB option). The contaminant library content of common MS contaminant sequences (e.g., trypsin, keratins, and albumin) was included in the database. The modifications selected in the search were carbamidomethyl (C), deamination (NQ, variable), and oxidation (M, variable). Enzyme trypsin (Fully-specific), 2 maximum missed cleavages and initial precursor mass tolerance of 10 ppm were set as Comet parameters. The acceptable FDR (False Discovery Rate) estimates of 3% at spectral, 2% at peptide, and 1% at protein levels; and MS and MS/MS tolerance errors of 10 ppm were added as parameters in the Filtering (SEPro) options.

The functional annotation of the identified proteins was performed using UniProtKB/Swiss-Prot database, and the graphical summary of functional classification was created using GO terms through the WEGO 2.0 tool (Web Gene Ontology Annotation Plot) (Ye et al., 2018).

2.10 Data availability

This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession JANVFA000000000. The version described in this paper is version JANVFA010000000.

3 Results

3.1 Protein profile, morphological, and toxicity analysis

Cry proteins profile of *Bt* TOL651 revealed the presence of two major protein bands of approximately 130 and 65 kDa size; indicating the presence of both Cry1 and Cry2 proteins, also observed in the reference strain (Fig. 1A). The morphological analysis revealed the presence of spores, bipyramidal and cuboidal crystals in this strain, also indicating the expression of Cry1(bipyramidal crystals) and Cry2 (cuboidal crystals) proteins (Fig. 1B).

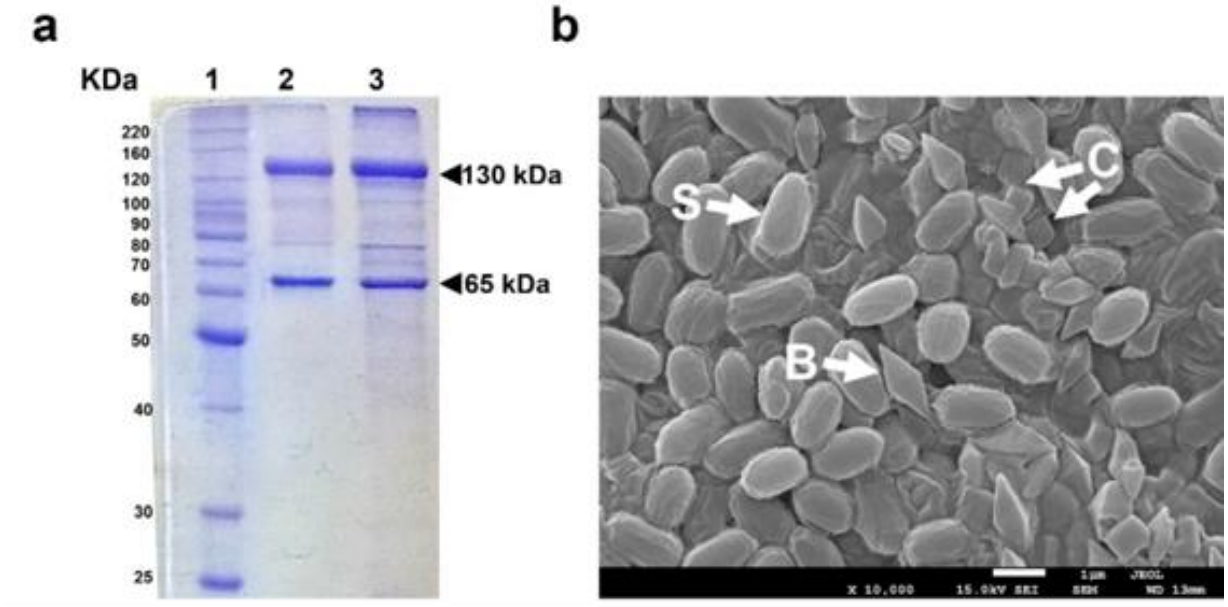

Fig. 1 SDS-PAGE analysis of Cry proteins and scanning electron microscopic of spore-crystal mixture of *Bt* TOL651. **a**- Protein profile: Lane 1 - molecular mass markers; Lane 2- HD-1; Lane 3- TOL651. Arrows indicate likely protein band size. **b**- The view of spores (S) and bipyramidal (B), and cuboidal (C) crystals enlarged 2,000 (x) times and approximated 10,000 times (x).

The insect bioassays using spore-crystal mixtures of *Bt* TOL651 and the reference *Bt* HD-1 strain showed that both strains were toxic to *D. saccharalis* and *A. gemmatalis* (Table 1). However, the *Bt* TOL651 showed significantly higher toxicity to *D. saccharalis* and *A. gemmatalis* when compared to the *Bt* HD-1 strain (Table 1). The RT_{50} estimate showed that the *Bt* TOL651 strain is 1.97- and 1.75-fold more toxic than the *Bt* HD-1 strain against *D. sacharalis*; and *A. gemmatalis*, respectively (Table 1).

Table 1 Lethal concentration of spore-crystal mixture from *Bt* TOL651 against larvae of *A. gemmatalis* and *D. saccharalis*.

D. saccharalis	**Strain**	**Slope ± SE**	**LC_{50}* ng.cm^{-2} (IC)^{c}**	**LC_{95}** ng.cm^{-2} (IC)^{c}**	**χ^2**	***P***	**RT_{50}**
	TOL651	2.69 ± 0.26	73.77 (63.23-87.56)	300.37 (222.34-461.86)	0.62	0.94	1.97
	HD-1	2.78 ± 0.30	145.11 (125.10-164.81)	566.51 (499.94-794.44)	5.28	0.15	
A. gemmatalis	TOL651	4.08 ± 0.48	1.45 (1.33-1.56)	3.66 (3.06-4.84)	0.51	0.91	1.75
	HD-1	3.99 ± 0.56	2.55 (2.30-2.79)	6.58 (5.40-9.18)	3.21	0.20	

*Lethal concentration
**Confidence interval 95% probability.
RT_{50} = Toxicity ratio = LC_{50} HD-1 / LC_{50}TOL651 (Robertson et al., 2017)

3.2 Genomic characterization

The draft genome of *Bt* TOL651 was obtained and consists of ~ 6.17 Mb with 35.3% GC content (Table 2). A total of 7003 coding sequences (CDS) were found, out of which 4812 proteins had functional assignments and 2191 were considered hypothetical proteins. Sixty-three tRNA and 5 rRNA genes were also annotated (Table 2).

Table 2 Genome features of *Bt* TOL651 strain.

General Features	**Value**
Scaffolds (no.)	256
N50 (bp)	819,816
L50 (no.)	3
N bases (%)	1.5
Completeness (%)	98
Genome Length (bp)	6,176,245
GC Content (%)	35.34
Coding sequences (CDS) (no.)	7003
tRNA (no.)	63
rRNA (no.)	9
Proteins with functional assignments (no.)	4812
Hypothetical proteins (no.)	2191

In the subsystem class distribution, most of the genes were involved, in decreasing order: amino acids and derivatives metabolism (387); carbohydrate (265); cofactors, vitamins, prosthetic groups, pigments metabolism subsystems (161); protein metabolism (155) and nucleosides/nucleotides metabolism (118) (Fig. 2).

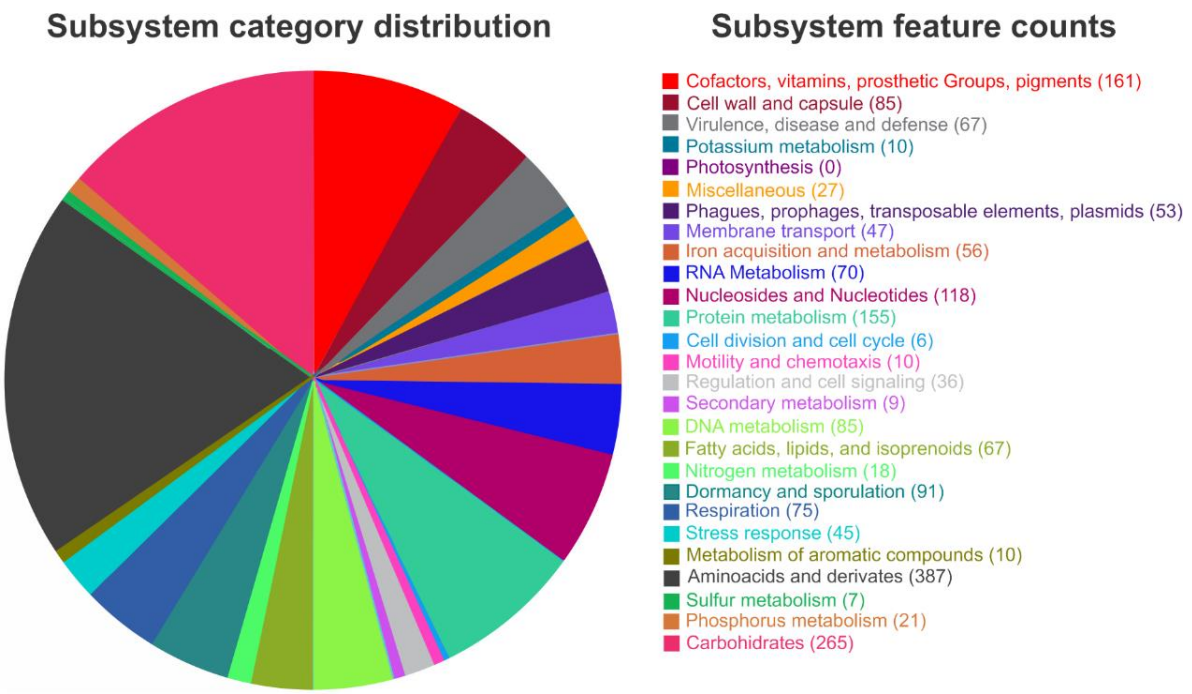

Fig. 2 Subsystem category distribution of genes in the genome of *Bt* TOL651 based on RAST annotation server.

3.3 *Bt* TOL651 relationships with other *Bt* strains

In the search for *Bt* genomes highly correlated to *Bt* TOL651, different *Bt* strains' genomes were compared with *Bt* TOL651. ANI values were obtained and showed that *Bt* TOL651 is highly similar to other *Bt* strains (ANI ≥ 94%). High ANI values were observed for *Bt* TOL651 and other *Bt* strains such as *Bt* kurstaki T03a001 (99.06 %), *Bt* kurstaki HD73 (99.00 %), *Bacillus* sp. G3 (98.88%), Bt YC-10 (98.77%), *Bt mexicanensis* 27 (98.63%), *Bt* NBIN-66 (98.41%), *Bt galleriae* HD-29 (98.28%) and *Bt aizawai* Leap01 (98.04%). However, *Bacillus cereus* (*Bc*) B4158 genome also was highly correlated with *Bt* TOL651 (ANI= 98.16 %) (Fig. 3).

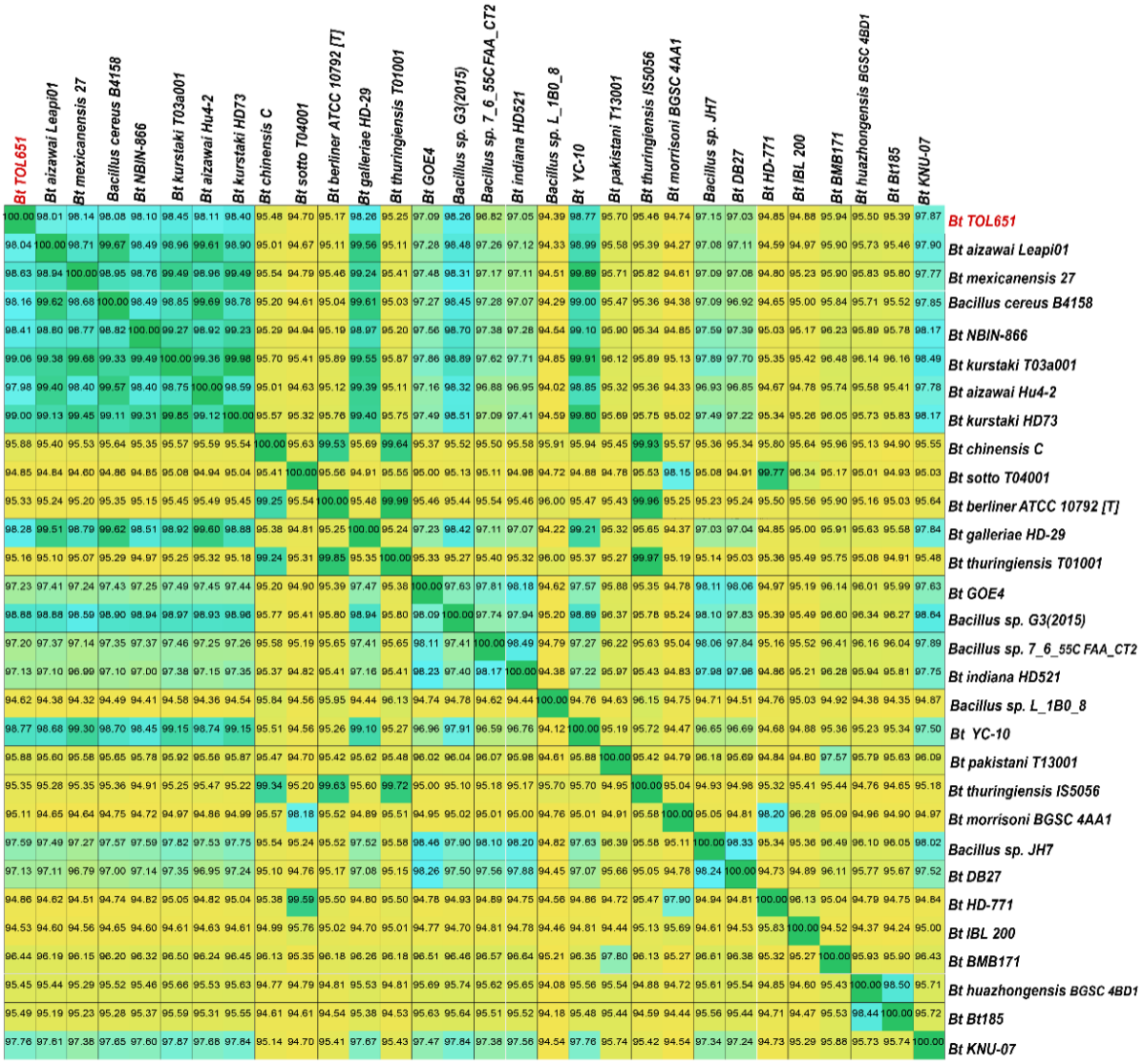

Fig. 3 Heat map of average nucleotide identity (ANI) based on genomic comparison *Bt* TOL651 and other 29 related strains. The percentages identities are listed on the map.

The phylogenetic analysis using gyr*B* gene indicates that *Bt* TOL651 was clustered with *Bt kenyae* BGSC-4F2 and IEBC-T04B001 strains (Fig. 4).

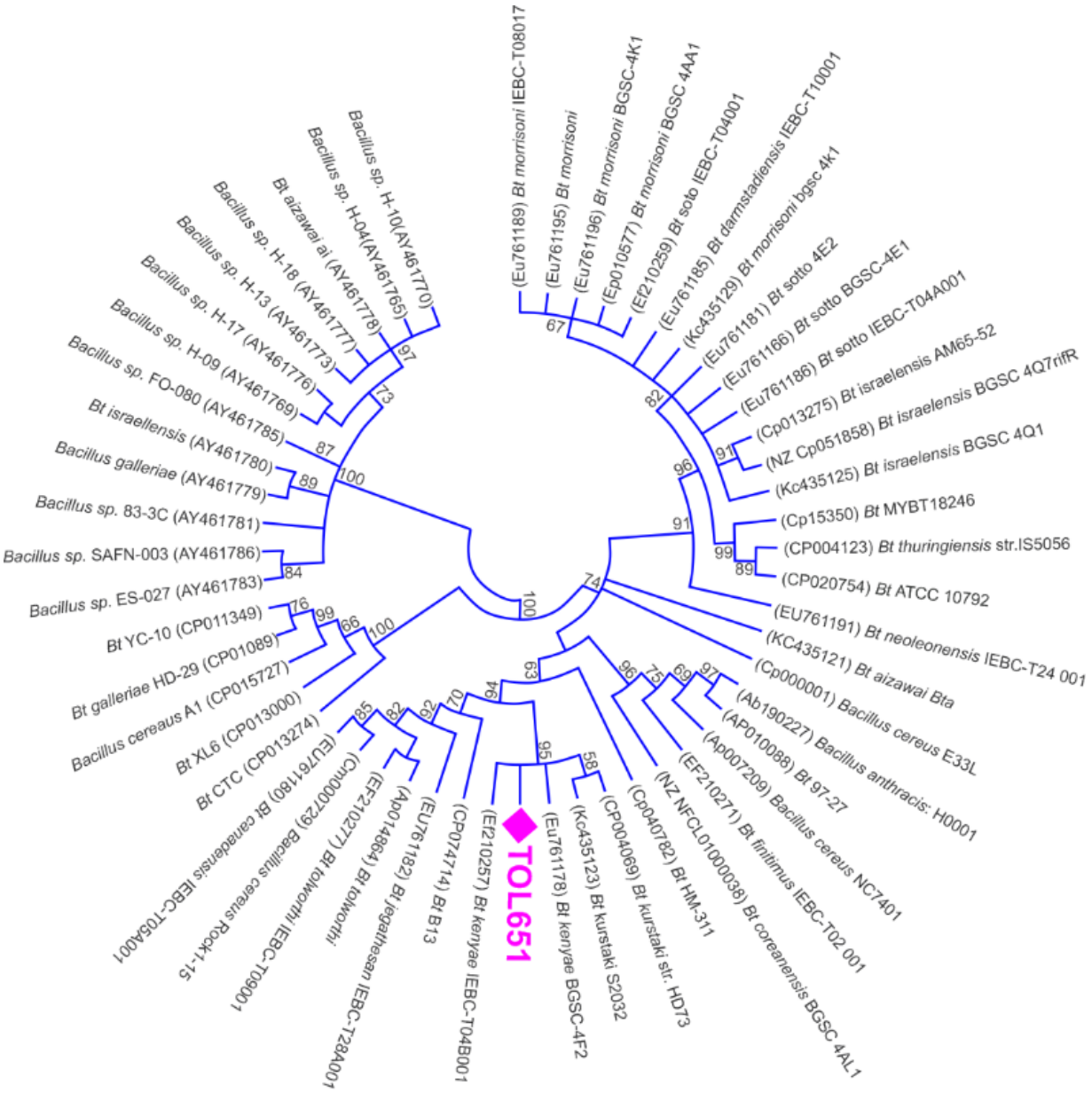

Fig. 4 Phylogenetic analysis using the Neighbor-Joining method. *gyrB* gene sequences of the *Bt* TOL651 and other *Bacillus* ssp. strains were used for the analysis. Bootstrap values (%) presented at the branches were calculated from 1,000 replications. Bootstrap values <50% were disregarded for branches in the graphical representation of the tree.

3.4 Genes related to pesticidal proteins, virulence-factors, and antibiotic resistance

Based on the *Bt* database and Btoxin_Digger, a total of four *cry1* and *cry2*-types genes were found in two scaffolds (18 and 21) in the genome of *Bt* TOL651. The CDS regions peg.5608, peg.5616, and peg.6113 showed high homology to the *cry2Aa9*, *cry1Ia44*, *and cryAc5* genes, respectively. On the other hand, the peg.5617 CDS region showed high homology to the *cry1Aa18* gene in the custom *Bt* database, but not was found in the Btoxin_Digger (Table 3, Fig. 5). Finally, the peg.3270 CDS region was retrieved using Btoxin_Digger and custom *Bt* database, demonstrating high homology to the *spp1Aa1* gene (~ 80%).

Table 3 Identification of genes coding pesticidal proteins-like in the *Bt* TOL651.

Sequence_id	CDS_id	Length (aa)	Hit_id	Coverage (%)	Identity (%)	E-value	Accession
Scaffold_5	peg.3270	513	Spp1Aa1[a,b]	99.21[a] 97.47[b]	80.48[a] 80.50[b]	0.0 0.0	BAF62176
Scaffold_18	peg.5608	634	Cry2Aa9[a,b]	100[a] 99.84[b]	100[a] 96.7[b]	0.0 0.0	ABR68091.1
Scaffold_18	peg.5616	720	Cry1Ia44[a,b]	100[a] 99.86[b]	100[a] 100[b]	0.0 0.0	QBO24620
Scaffold_18	peg.5617	521	Cry1Aa18[b]	98.66[a]	100[a]	0.0	AEH31438
Scaffold_21	peg.6113	1178	Cry1Ac5[a,b]	100[a] 99.92[b]	99.92[a] 98.60[b]	0.0 0.0	AAA22339

[a]Btoxin_Digger
[b]Customized *Bt* database

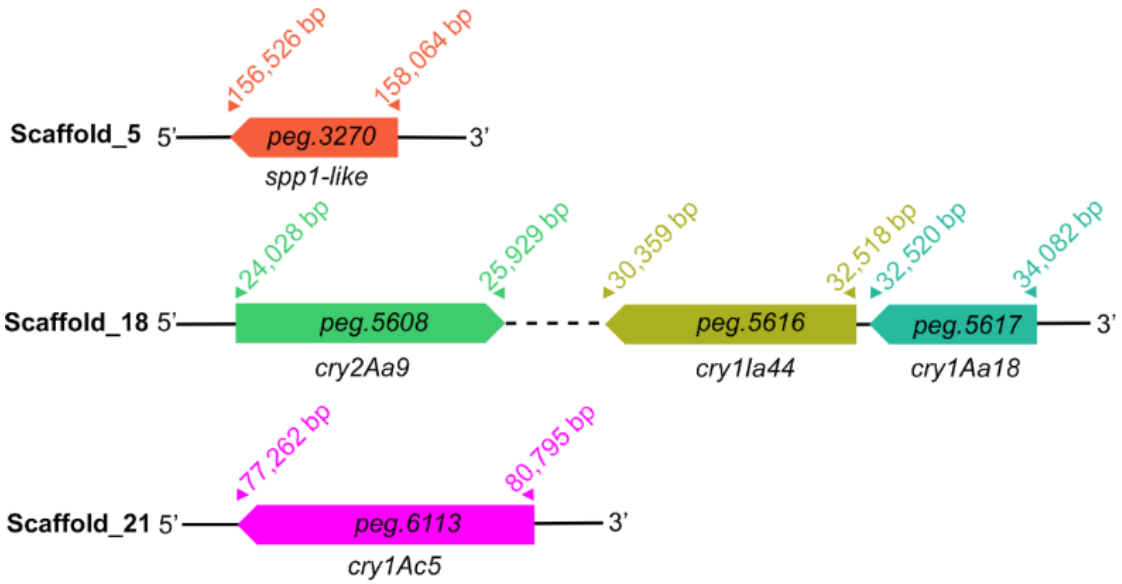

Fig. 5 Representation of the genome position of pesticidal protein-like identified in the scaffolds of the *Bt* TOL651 strain.

Virulence factor genes from different classes (adherence, enzyme, immune evasion, iron acquisition, regulation, secretion system, toxins, acid resistance, magnesium uptake, surface protein anchoring, and others) were also found in the genome sequence of *Bt* TOL651 (Table S1). Among these genes we can point out enzymes such as phospholipases (*plcA*, *piplc*) and metalloproteases (*inhA*); and other toxins such as hemolysins (*cytK*, *hlyIII*, *hblA*, *hblC*, *hblD*) and enterotoxins (*nheA*, *nheB*, *nheC*). (Table S1). Four putative antibiotic resistance genes were identified in the genome of *Bt* TOL651, including *BcII* gene (subclass B1 beta-lactamase) (% identity: 90.12; % length of reference sequence: 122.66); *BcI* and two *BcIII* genes (class A beta-lactamase) (% identities: 95.42, 86.83 and 76.58, respectively, % length of reference sequences: 100.65; 100.32 and 100, respectively).

3.5 Proteomic of spores-crystal mixture

The general functional classification carried out by LC-MS/MS analysis of the spore-crystal mixture of *Bt* TOL651 revealed that detected proteins sequences were involved in 11 GO terms related to cellular components, 10 GO terms related to molecular functions, and 19 terms related to biological processes (Fig. 6). In the cellular component groups, most proteins were related mainly to cell and membrane components. Furthermore, the molecular function classification was represented by proteins with catalytic and binding activities; next, in the biological process category, most proteins belonged to metabolic and cellular processes.

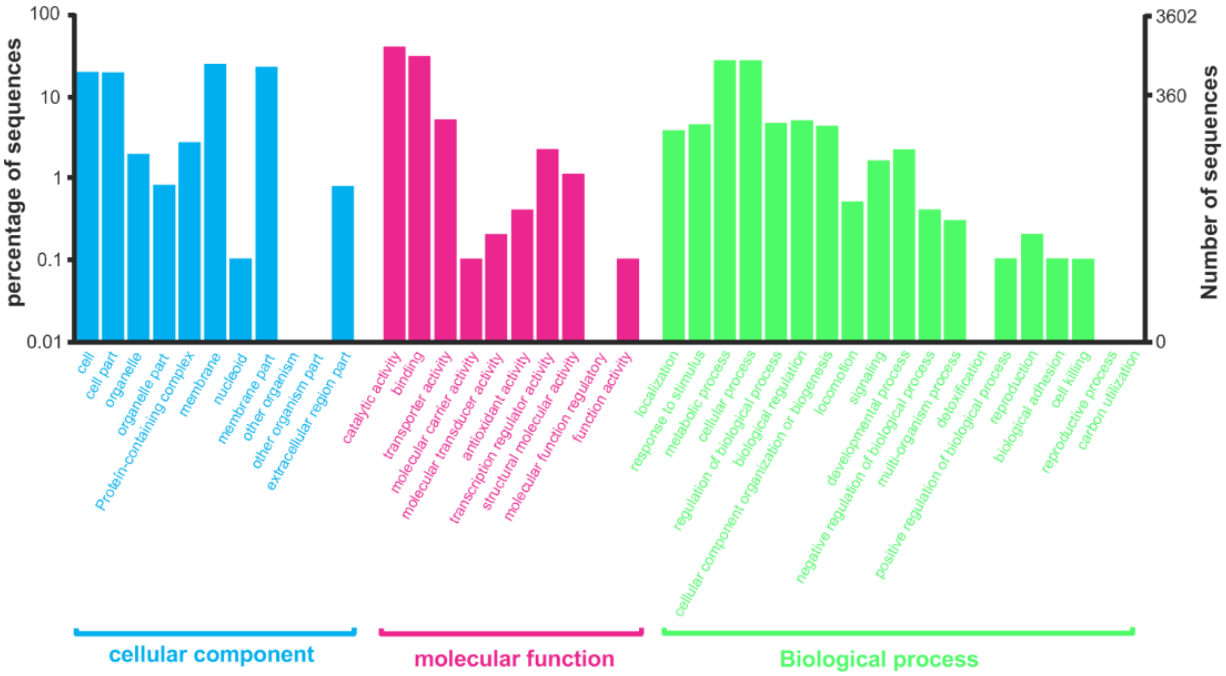

Fig. 6 Functional annotation and classification for LC–MS/MS-identified proteins of TOL651 in the spore-crystal mixture.

The comparison among genomic and proteomic data indicated that 24 coding sequences were identified in the proteins' sequences and based on unique peptide count, the most abundant proteins were: Cry1Ac5, Cry2Aa9, and Cry1Aa18 (Table 4) (Table S2). Except for Cry1Ia44 (peg.5616) which had no unique peptide, all Cry proteins identified in the genome were expressed in the sporulation phase (72 h of incubation). The Cry1Ac5 (peg.6113) showed the highest number of peptides detected, with 85 in total and 82 unique sequences (Table 4) (Table S2). Additionally, regarding the pesticidal proteins, the Inhibitor A metalloprotease (InhA1), a potential pathogenic factor, was also found in the proteome of *Bt* TOL651 (Table 4) (Table S2). Other spore-associated proteins (spore coat and forespore-specific proteins), peptides associated with protein metabolism (Elongation factor Tu, Shock protein Hsp20 family, and Chaperone protein DnaK), and other functions (Aminopeptidase, Glycerophosphoryl phosphodiesterase, Choline binding protein, DUF3915 domain-containing, Acid endopeptidase, Neutral protease B and DNA-binding protein) were also detected (Table 4).

Table 4 Identification of pesticidal and other proteins in the spore/crystal mixture of the *Bt* TOL651 strain detected by LC-MS/MS analysis.

CDS id	Description[a]	Length (bp)	Peptide sequence (No.)	Unique Peptides (No.)[d]	Coverage[e]	Protein Score[f]	NSAF[g]
peg.6113	Cry1Ac5[b]	1177	85	82	0.5582	311.262	0.3234803
peg.5608	Cry2Aa9[b]	633	46	46	0.5987	149.316	0.1884978
peg.5617	Cry1Aa18[b]	520	27	24	0.5635	101.849	0.2065138
peg.5616	Cry1Ia44	750	2	0	0.0167	6.19	0.0489361
peg.4938	Inhibitor A metalloprotease (InhA1)	796	8	8	0.0804	21.097	0.0054508
peg.4935	Spore coat-associated protein 1	197	6	6	0.1777	16.659	0.0330371
peg.6532	Spore coat protein CotG	179	5	5	0.1229	19.754	0.0212096
peg.1595	Glycerophosphoryl phosphodiesterase	314	5	5	0.1783	14.459	0.0086363
peg.3277	Choline binding protein (PcpA)[c]	310	5	5	0.1548	16.069	0.0122468
peg.4409	Spore coat protein (CotB)	169	5	5	0.2781	15.245	0.0192554
peg.2864	Shock protein, Hsp20 family	154	4	4	0.2597	10.137	0.0140873
peg.920	Hypothetical protein	247	4	4	0.1093	12.996	0.0109789
peg.969	DUF3915 domain-containing[c]	122	4	4	0.4016	12.488	0.0222278
peg.570	Forespore-specific protein	213	4	4	0.1549	10.307	0.0101852
peg.3555	Aminopeptidase	466	4	4	0.1438	13.59	0.0046554
peg.4408	Spore coat protein CotB	149	4	4	0.4966	13.189	0.0182
peg.4933	Spore coat-associated protein 2	195	4	4	0.2	10.628	0.016688
peg.688	Chaperone protein DnaK	611	3	3	0.0426	10.54	0.0035506
peg.1529	Uncharacterized protein (YxeE)	109	3	3	0.2018	6.706	0.0199031
peg.3676	DNA-binding protein	170	2	2	0.1882	4.262	0.0063807
peg.2225	Acid endopeptidase	333	2	2	0.0961	6.225	0.0032574
peg.3226	Neutral protease B (NprB)	591	2	2	0.0592	7.248	0.0036708
peg.65	Hypothetical protein	155	2	2	0.2065	6.864	0.0139964
peg.5901	Elongation factor Tu	395	2	2	0.1038	8.342	0.0041192

The main pesticidal proteins and the virulent factor are shown in bold.
[a]Annotation based on RASTtk
[b]Classification based on Btoxin_Digger and/or Customized *Bt* database
[c]Descrition based on BLASTx
[d]The number of peptide sequences that are unique to protein.
[e]The percentage of the protein sequence covered by identified peptides.
[f]The sum of the ion scores of all peptides that were identified.
[g]Normalized Spectral Abundance Factor, calculated using the number of spectra divided by the protein length and then normalized over the total of spectral counts/length for all the proteins in the sample.

4 Discussion

Bt TOL651 strain analyses presented a Cry proteins profile in SDS-PAGE gel with two major protein bands of approximately 130 and 65 kDa in size which is also associated with Cry1 and Cry2 proteins, respectively (Ganesh et al., 2018; Monnerat et al., 2007; Schnepf et al., 1998; Singh et al., 2021). In concordance with the protein profile, the crystals morphology of *Bt* TOL651 revealed bipyramidal forms associated with Cry1 proteins and cuboids formed by Cry2 proteins similar to HD-1 strain (Monnerat et al., 2007; Schnepf et al., 1998).

The TOL651 genome similarity analysis confirmed that this isolate is a *Bt* species since ANI values ≥ 95% (98-99%) were assigned in comparison with other genomes of this species (Richter and Rosselló-Móra, 2009). However, a high ANI value between *Bt* TOL651 and the *B. cereus* (*Bc* B4158) was also found, which corroborates the complex separation between *Bt* and *B. cereus* at the genomic level (Helgason et al., 2000; Lechuga et al., 2020; Zhou et al., 2022). Although, in the subspecies classification the phylogenetic study showed that the *Bt* TOL651 strain might have a close relationship with *Bt kenyae* strains.

Bt strains identified as *kenyae* subspecies have been reported to harbor c*ry1Ab*, *cry1Ac*, *cry1E*, and *cry2Aa* genes and showed toxic activity against lepidopteran insects (Chang et al. 1999; Hire et al. 2008, 2009). *Bt* TOL651 harbors *cry2Aa9*, *cry1Ia44* and *cryAc5* genes. HD-1 strain, a *kurstaki* subspecies, harbors *cry1Aa*, *cry1Ab*, *cry1Ac*, *cry1Ia*, *cry2Aa*, *cry2Ab*, and *vip3Aa* genes of pesticidal proteins (Zhu et al., 2015).

In addition, *Bt* TOL651 harbored enterotoxins such as hemolysins genes (Hemolysins I, III, and IV) presented in other species of *Bacillus cereus* sensu lato group and detected in new *Bt* strains (Kim et al., 2015; Ma et al., 2020; Lechuga et al., 2020). Although *Bt* is considered a safe bioinsecticide for non-target organisms (Pohare et al., 2021), the presence of these genes in *Bt* TOL651 requires their validation for safety use in crops, as well this toxicity for humans, since we are proposing its use as a spore-crystal mixture. Considering the potential antibiotic resistance of *Bt* TOL651, the putative antibiotic resistance genes were screened in the genome sequence indicating the presence of genes resistant to beta-lactam antibiotics, corroborating other studies (Luna et al., 2007; Kaze et al., 2021). Banik et al. (2019) have demonstrated the sensitivity of *Bt* strains against clinically important antibiotics, indicating it is a safe biocontrol agent for crop application without any harm to consumers. However, additional experimental antibiogram tests need to be conducted to validate the resistance or sensitivity of the TOL651 to antibiotics.

The combined genomic and proteomic analysis described in this work revealed that the spore-crystal mixture of the *Bt* TOL651 has a high proportion of the pesticidal proteins of Cry1Ac5, Cry2Aa9, and Cry1Aa18. These Cry proteins showed in other studies to be toxic to lepidopteran insect pests (Dammak et al., 2015; dos Santos et al., 2009; Sun et al., 2022). Similarly, genomic and proteomic of the *Bt* 4.0718 strain indicated expression of Cry2Aa, Cry1Aa, and Cry1Ac (Rang et al., 2015; Huang et al., 2012). The *Bt* HD-1 strain from Caballero et al. (2020) using the same omics approach indicated the expression of Cry1Aa, Cry1Ac, Cry1Ab, and Cry2Aa.

Although present in the genome of *Bt* TOL651, the Cry1Ia44 protein was not detected in the spore-crystal mixture by proteomic analysis. Despite harboring the *cry1Ia* gene sequence, in other *Bt* strains the expression of this protein was also not detected, based on LC-MS/MS approach (Huang et al., 2012; Quan et al., 2016). This result, corroborating with other studies, has suggested the cryptic nature of the *cry1I* genes based on their absence in parasporal crystals (Gleave et al., 1993; Tailor et al., 1992). However, the genomic and proteomic study of the *Bt* AB1 strain revealed three unique peptides of Cry1Ia, their expression being considered at a trace level (Baragamaarachchi et al., 2019).

The Cry1Ac5 was the most abundant spore-crystal of *Bt* TOL651. The gene coding Cry1Ac5 has been described in *Bt* isolates from Warehouses in China (Hongyu et al., 2000) and in *Bt* strain isolated from India, toxic to *Spodoptera frugiperda* (Sathyan et al., 2022). In contrast to TOL651, another wild-type *Bt* isolates, and commercial HD-1 strain expressed Cry2Aa in greater proportion (Huang et al., 2012; Caballero et al. 2020) indicating that the abundance of Cry-like proteins may change according to the strain.

Bioassays with the spore-crystal mixture showed, in comparison with the commercial strain *Bt* HD-1, that *Bt* TO651 was more toxic for both insects with *A. gemmatalis* the most susceptible. Studies showed the toxicity level of each Cry protein varies with the insect species. For example, Bel et al. (2017) found that Cry1Ac was most toxic than Cry1Aa against *A. gemmatalis*, but the opposite occurred when these proteins were tested on *Chrysodeixis includens* (Walker, 1858) (Noctuidae: Lepidoptera). *D. saccharalis* was susceptible to Cry1Aa and Cry1Ac, of which Cry1Ac was more effective (Davolos et al., 2015; Tan et al., 2011). However, this insect was more susceptible to a mixture of Cry1 and Cry2 proteins than tested separately (Macedo et al., 2012). This can happen because Cry1Ac and Cry2Aa proteins do not share the same midgut receptor binding sites; therefore, a synergic effect can be expected when both proteins are present (Macedo et al., 2012).

Cry1Aa, Cry1Ab, and Cry1Ac proteins share binding sites (Davolos et al., 2015). In contrast to HD-1 (Caballero et al., 2020), TOL651 expressed only Cry1Aa18 and Cry1Ac5. This findings has suggested that Cry1Ac has a high affinity in comparison to Cry1Aa for the shared binding site due to divergences in domain II of proteins (Hernández-Rodríguez et al., 2013). So, the presence of Cry1Aa18 and Cry1Ac5 in the spore-crystal mixture of TOL651 may not affect the action of Cry1Ac5. Thus, our results suggest that a major proportion of Cry1Ac5 protein in the spore-crystal mixture along Cry1Aa18 and Cry2Aa9 proportions may explain the toxicity of *Bt* TOL651 against *A. gemmatalis* and *D. saccharalis*.

In addition to Cry proteins, the Inhibitor A metalloprotease (InhA1) was also detected in the spore-crystal mixture of *Bt* TOL651. InhA1 produced early in the sporulation phase allows the neutralization of the host immune system by specifically degrading the insect attacin and cecropin proteins (Miyoshi and Shinoda, 2000; Pohare et al., 2021). Besides, Dammak et al. (2015) mentioned that InhA1 within a spore-crystal mixture can enhance the pathogenic effect of Cry1-Cry2 proteins since can cause disorganization of the intestinal epithelium and delay a possible resistance caused by the intensive use of Cry proteins. Although InhA1 expression has been reported in spore-mixture of *Bt* strains (Banik et al., 2019; Khorramnejad et al., 2020), the researches have indicated the absence of InhA1 in the late sporulation phase (Li et al., 2012), and different levels in transcription and expression of InhA1 between *Bt* strains, suggesting a possible strategy to adapt to various hosts (Zhu el al., 2015). The deficiency in the expression of the virulence factor camelysin also produced for *Bt* could be involved in the lack of expression of InhA (Yin et al., 2015).

Neutral protease B (NprB) (also named NprA and Npr99) was also present in the spore-crystal mixture of *Bt* TOL651 and associated with virulence of *Bacillus cereus*, degrading host tissues resulting in increasing tissue permeability to the pathogen (Chung MC et al., 2006). Heat shock protein Hsp20 and the Elongation factor Tu were also detected in the spore-crystal mixture of *Bt* TOL651, and are necessary for the formation of crystals in *Bt* strains (Ding et al., 2009). The Hsp20 protein supports other proteins in refolding and preventing protein degradation (Xie et al., 2019).

Biopesticides are commonly used in multi-strain consortia and represent better cost-effectiveness than constructing recombinant or purified toxins for the development of products (Sreshty, Kumar and Murty, 2011). Synergism between different crystalline proteins produced by two *Bt* strains that do not compete for the same binding site has shown enhanced activity against lepidopteran pests (Konecka et al., 2012). The genes and proteins identified in

the genome-proteomic step study of *Bt* TOL651 will experimentally facilitate the determination of the potential of synergism between TOL651 and other strains.

In conclusion, our findings showed the potential use of the Brazilian *Bt* TOL651 strain in the control of *D. saccharalis* and *A. germmatalis*, of which *A. germmatalis* was most susceptible. *Bt* TOL651 was closely related to *kenyae* subspecies and expressed mainly Cry1Aa18, Cry1Ac5, and Cry2Aa9 pesticidal proteins in the spore-crystal mixture, with Cry1Ac5 being the most abundant protein. The virulence factor InhA1 may contribute to the pathogenicity of the *Bt* TOL651. The genomic-proteomic approach used in this study allowed a better understanding of *Bt* TOL651 pathogenicity, representing an important step for the development and monitoring of potential new bioinsecticides.

5 References

Alcock, BP, Huynh, W, Chalil, R, Smith, KW, Raphenya, AR, Wlodarski, MA, McArthur, AG (2023) CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance DatabaseNucleic Acids Res., 51: D690-D699.

Andrews S (2015) FastQC: a quality-control tool for high-throughput sequence. Retrieved Octuber from https://www.bioinformatics.babraham.ac.uk/projects/fastqc Accessed 18 June 2020

Arthur S, Dara, SK (2019) Microbial biopesticides for invertebrate pests and their markets in the United States. J. Invertebr. Pathol. 165:13-21. https://doi.org/https://doi.org/10.1016/j.jip.2018.01.008

Banik, A, Chattopadhyay, A, Ganguly, S., Mukhopadhyay, SK (2019) Characterization of a tea pest specific *Bacillus thuringiensis* and identification of its toxin by MALDI-TOF mass spectrometry. Ind Crops Prod, 137, 549-556. https://doi.org/10.1016/j.indcrop.2019.05.051

Bankevich A, Sergey Nurk, Dmitry Antipov, Alexey A Gurevich, Mikhail Dvorkin, Alexander S Kulikov, Valery M Lesin, Sergey I Nikolenko, Son Pham, Andrey D Prjibelski, Alexey V Pyshkin, Alexander V Sirotkin, Nikolay Vyahhi, Glenn Tesler, Max A Alekseyev, Pevzner, PA (2012) SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19:455-477. https://doi.org/https://doi.org/10.1089/cmb.2012.0021

Baragamaarachchi RY, Samarasekera JK, Weerasena OV, Lamour K, Jurat-Fuentes JL (2019) Identification of a native *Bacillus thuringiensis* strain from Sri Lanka active against Dipel-resistant *Plutella xylostella*. PeerJ 7:e7535. https://doi.org/https://doi.org/10.7717/peerj.7535

Bel Y, Sheets JJ, Tan SY, Narva KE, Escriche B (2017) Toxicity and Binding Studies of *Bacillus thuringiensis* Cry1Ac, Cry1F, Cry1C, and Cry2A Proteins in the Soybean Pests *Anticarsia gemmatalis* and *Chrysodeixis* (Pseudoplusia) *includens*. Appl. Environ. Microbiol. 83:e00326-00317. https://doi.org/https://doi.org/10.1128/AEM.00326-17

Bosi E, Donati B, Galardini M, Brunetti S, Sagot MF, Lió P, Crescenzi P, Fani R, Fondi M (2015) MeDuSa: a multi-draft based scaffolder. Bioinformatics 31:2443-2451. https://doi.org/https://doi.org/10.1093/bioinformatics/btv171

Bravo A, Gill SS, Soberón M (2007) Mode of action of *Bacillus thuringiensis* Cry and Cyt toxins and their potential for insect control. Toxicon 49:423-435. https://doi.org/https://doi.org/10.1016/j.toxicon.2006.11.022

Bravo A, Likitvivatanavong S, Gill SS, Soberón M (2011) *Bacillus thuringiensis*: A story of a successful bioinsecticide. Insect Biochem. Mol. Biol. 41:423-431. https://doi.org/https://doi.org/10.1016/j.ibmb.2011.02.006

Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch, GD, Shukla M, Thomason JA, Stevens R, Vonstein V, Wattam AR, Xia F (2015) RASTtk: A modular and extensible implementation of the RAST algorithm for building

custom annotation pipelines and annotating batches of genomes. Sci. Rep. 5:8365. https://doi.org/https://doi.org/10.1038/srep08365

Caballero J, Jiménez-Moreno N, Orera I, Williams T, Fernández AB, Villanueva M, Ferré J, Caballero P, Ancín-Azpilicueta C (2020) Unraveling the composition of insecticidal crystal proteins in *Bacillus thuringiensis*: A proteomics approach. Appl. Environ. Microbiol. 86:e00476-20. https://doi.org/10.1128/AEM.00476-20

Cardoso P, Fazion F, Perchat S, Buisson C, Vilas-Bôas G, Lereclus D (2020) Rap-Phr systems from plasmids pAW63 and pHT8-1 act together to regulate sporulation in the *Bacillus thuringiensis* serovar *kurstaki* HD73 strain. Appl. Environ. Microbiol. 86: e01238-01220. https://doi.org/https://doi.org/10.1128/AEM.01238-20

Castro BMdC, Martinez LC, Barbosa SG, Serrão JE, Wilcken CF, Soares MA, da Silva, AA, de Carvalho AG, Zanuncio JC (2019) Toxicity and cytopathology mediated by*Bacillus thuringiensis* in the midgut of *Anticarsia gemmatalis* (Lepidoptera: Noctuidae). Sci. Rep. 9:6667. https://doi.org/https://doi.org/10.1038/s41598-019-430740

Cerqueira, FB, Alves, GB, Corrêa, RFT, Martins, ES, Barbosa, LCB, do NASCIMENTO, IR, Aguiar, RDS (2016) Selection and characterization of *Bacillus thuringiensis* isolates with a high insecticidal activity against Spodoptera frugiperda (Lepidoptera: Noctuidae). Biosci.J 32:1522-1536

Chang JH, Je YH, Roh JY, Park HW, Jin BR, Lee DW, Kim S-H, Yang-W, Kang SK (1999) Isolation and characterization of a strain of *Bacillus thuringiensis* serovar *kenyae* encoding olny δ-endotoxin Cry1E. Appl. Entomol. Zool., 34:379-382. https://doi.org/10.1303/aez.34.379

Chung MC, Popova TG, Millis BA, Mukherjee DV, Zhou W, Liotta LA, Petricoin EF, Chandhoke V, Bailey C, SG, P (2006) Secreted neutral metalloproteases of *Bacillus anthracis* as candidate pathogenic factors. J. Biol. Chem. 281:31408-31418. https://doi.org/10.1016/S0021-9258(19)84053-X

da Silva SM, Silva-Werneck JO, Falcão R, Gomes AC, Fragoso RR, Quezado MT, Neto OB, Aguiar JB de Sá, MF, Bravo A, Monnerat RG (2004) Characterization of novel Brazilian *Bacillus thuringiensis* strains active against *Spodoptera frugiperda* and other insect pests. J. Appl. Entomol. 128:102-107. https://doi.org/https://doi.org/10.1046/j.1439-0418.2003.00812.x

Dammak I, Dammak M, Tounsi S (2015) Histopathological and combinatorial effects of the metalloprotease InhA1 and Cry proteins of *Bacillus thuringiensis* against *Spodoptera littoralis*. Int. J. Biol. Macromol. 81:759-762. https://doi.org/https://doi.org/10.1016/j.ijbiomac.2015.09.006

Daquila BV, Scudeler EL, Dossi FC, Moreira DR, Pamphile JA, Conte H (2019) Action of *Bacillus thuringiensis* (Bacillales: Bacillaceae) in the midgut of the sugarcane borer *Diatraea saccharalis* (Fabricius, 1794) (Lepidoptera: Crambidae). Ecotoxicol. Environ. Saf. 184:109642. https://doi.org/https://doi.org/10.1016/j.ecoenv.2019.109642

Davolos CC, Hernández-Martinez P, Crialesi-Legori PC, Desidério JA, Ferré J, Escriche B, Lemos MV (2015). Binding analysis of *Bacillus thuringiensis* Cry1 proteins in the sugarcane borer, *Diatraea saccharalis* (Lepidoptera: Crambidae). J. Invertebr. Pathol. 127:32-34. https://doi.org/https://doi.org/10.1016/j.jip.2015.01.013

Day M, Ibrahim M, Dyer D, Bulla L (2014) Genome Sequence of *Bacillus thuringiensis* subsp. *kurstaki* Strain HD-1. Genome. Announc. 2:e00613-00614. https://doi.org/https://doi.org/10.1128/genomeA.00613-14

de Matos M, Santos F, Eichler P (2020) Sugarcane world scenario. In: Sugarcane biorefinery, technology and perspectives. Elsevier Inc., pp. 1–19 https://doi.org/10.1016/B978-0-12-814236-3.00001-9

de Oliveira WS, Sakuno CIR, Miraldo LL, Tavares MA, Komada KM, Teresani D, Santos JL, Huang F (2022) Varied frequencies of resistance alleles to Cry1Ab and Cry1Ac among Brazilian populations of the sugarcane borer, *Diatraea saccharalis* (F.). Pest. Manage. Sci. n/a, 1;14. https://doi.org/https://doi.org/10.1002/ps.7133

da Silva, I. H. S., de Freitas, M. M., Polanczyk, R. A. (2022) *Bacillus thuringiensis*, a remarkable biopesticide: from lab to the field. In Biopesticides (pp. 117-131). Woodhead Publishing. https://doi.org/10.1016/B978-0-12-823355-9.00021-3

Dinardo-Miranda, LL, Fracasso, JV, Da Costa, VP, Dos Anjos, IA, Lopes, DOP (2013) Reação de cultivares de cana-de-açúcar à broca do colmo. Bragantia, Campinas, 72:29-34.

Ding X, Huang J, Xia L, Li X, Yuan C, Dan S (2009) A proteomic analysis approach to study insecticidal crystal proteins from different strains of *Bacillus thuringiensis*. Biocontrol. Sci. Technol. 19:289-299. https://doi.org/https://doi.org/10.1080/09583150902749984

dos Santos KB, Neves P, Meneguim AM, dos Santos RB, dos Santos WJ, Boas GV, Dumas V, Martins E, Praça LB, Queiroz P, Berry C, Monnerat R (2009) Selection and characterization of the *Bacillus thuringiensis* strains toxic to *Spodoptera eridania* (Cramer), *Spodoptera cosmioides* (Walker) and *Spodoptera frugiperda* (Smith) (Lepidoptera: Noctuidae). Biol. Control 50:157-163. https://doi.org/https://doi.org/10.1016/j.biocontrol.2009.03.014

Dulmage, HTB (1973) Thuringiensis U. S. Assay Standard: Report on the Adoption of a Primary U.S. Reference Standard for Assay of Formulations Containing the delta-Endotoxin of *Bacillus thuringiensis*. Bull. Entomol. Soc. Am. 19:200–202.

Finney DJ (1971) Probit analysis, Cambridge University Press. Cambridge, UK.

Frankenhuyzen, KV (2009) Insecticidal activity of *Bacillus thuringiensis* crystal proteins. J. Invertebr. Pathol. 101: 1-16. http://dx.doi.org/10.1016/j.jip.2009.02.009.

Frankland, GC, Frankland, PF (1887) XI Studies on some new micro-organisms obtained from air. Philosophical Transactions of the Royal Society of London. (B.), 178:257-287. https://doi.org/10.1098/rstb.1887.0011

Ganesh KN, Reyaz A, Balakrishnan N (2018) Molecular characterization of an indigenous lepidopteran toxic *Bacillus thuringiensis* strain T532. J. Biol. Control 32:246-251. https://doi.org/https://doi.org/10.18311/jbc/2018/21604

Gleave AP, Williams R, Hedges RJ (1993) Screening by polymerase chain reaction of *Bacillus thuringiensis* serotypes for the presence of cryV-like insecticidal protein genes and characterization of a cryV gene cloned from *B. thuringiensis* subsp. *kurstaki*. Appl. Environ. Microbiol. 59:1683-1687. https://doi.org/https://doi.org/10.1128/aem.59.5.1683-1687.1993

Gomis-Cebolla J, Scaramal Ricietto AP, Ferré J (2018). A genomic and proteomic approach to identify and quantify the expressed *Bacillus thuringiensis* proteins in the supernatant and parasporal crystal. Toxins (Basel) 10:193. https://doi.org/https://doi.org/10.3390/toxins10050193

Greene GL, Leppla NC, Dickerson WA (1976) Velvetbean Caterpillar: A Rearing Procedure and Artificial Medium. J. Econ. Entomol. 69:487-488. https://doi.org/10.1093/jee/69.4.487

Gurevich A, Saveliev V, Vyahhi N, Tesler G (2013) QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072-1075. https://doi.org/https://doi.org/10.1093/bioinformatics/btt086

Heinrichs R, Otto R, Magalhães A, Meirelles GC (2017). Importance of Sugarcane in Brazilian and World Bioeconomy, in: Dabbert S, Lewandowski I, Weiss J, Pyka A (Eds.), Knowledge-Driven Developments in the Bioeconomy: Technological and Economic Perspectives. Springer International Publishing, Cham, pp. 205-217. https://doi.org/10.1007/978-3-319-58374-7_11

Helgason E, Økstad OA, Caugant DA, Johansen HA, Fouet A, Mock M, Hegna I, Kolstø A (2000) *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis*; one Ssecies on the basis of genetic evidence. Appl. Environ. Microbiol. 66:2627-2630. https://doi.org/https://doi.org/10.1128/AEM.66.6.2627-2630.2000

Hensley SD, Hammond AM Jr (1968) Laboratory techniques for rearing the sugarcane borer on an artificial diet. J. Econ. Entomol. 61:1742-1743. https://doi.org/https://doi.org/10.1093/jee/61.6.1742

Hernández-Rodríguez, CS, Hernández-Martínez, P, Van Rie, J, Escriche, B, Ferré, J (2013) Shared midgut binding sites for Cry1A. 105, Cry1Aa, Cry1Ab, Cry1Ac and Cry1Fa proteins from *Bacillus thuringiensis* in two important corn pests, *Ostrinia nubilalis* and *Spodoptera frugiperda*. PloS one, 8:e68164. https://doi.org/10.1371/journal.pone.0068164

Hire RS, Makde RD, Dongre TK, D'souza SF (2008) Characterization of the cry1Ac17 gene from an indigenous strain of *Bacillus thuringiensis* subsp. *kenyae*. Curr. Microbiol. 57:570-574. https://doi.org/10.1007/s00284-008-9244-3

Hire RS, Makde RD, Dongre TK, D'souza SF (2009) Expression, purification and characterization of the Cry2Aa14 toxin from *Bacillus thuringiensis* subsp. kenyae. Toxicon. 54: 519-524. https://doi.org/10.1016/j.toxicon.2009.05.022

Hongyu, Z, Ziniu, Y, Wangxi, D (2000) Composition and ecological distribution of Cry proteins and their genotypes of *Bacillus thuringiensis* isolates from warehouses in China. J. Invertebr. Pathol., 76:191-197. https://doi.org/10.1006/jipa.2000.4970

Horikoshi RJ, Ferrari G, Dourado PM, Climaco JI, Vertuan HV, Evans A, Pleau M, Morrell K, José MO, Anderson H, Martinelli S, Ovejero RF, Berger GU, Head G (2022) MON 95379 *Bt* maize as a new tool to manage sugarcane borer (*Diatraea saccharalis*) in South America. Pest. Manage. Sci. 78:3456-3466. https://doi.org/https://doi.org/10.1002/ps.6986

Huang F, Chen M, Gowda A, Clark TL, McNulty BC, Yang F, Niu Y (2015) Identification, inheritance, and fitness costs of Cry2Ab2 resistance in a field-derived population of sugarcane borer, *Diatraea saccharalis* (F.) (Lepidoptera: Crambidae). J. Invertebr. Pathol. 130:116-123. https://doi.org/https://doi.org/10.1016/j.jip.2015.07.007

Huang S, Ding X, Sun Y, Yang Q, Xiao X, Cao Z, Xia L (2012) Proteomic analysis of *Bacillus thuringiensis* at different growth phases by using an automated online two-dimensional liquid chromatography-tandem mass spectrometry strategy. Appl. Environ. Microbiol. 78:5270-5279. https://doi.org/https://doi.org/10.1128/AEM.00424-12

Jeong H, Choi SK, Park SH (2017) Genome Sequences of *Bacillus thuringiensis* Serovar kurstaki Strain BP865 and *B. thuringiensis*, serovar *aizawai* Strain HD-133. Genome Announ. 5:e01544-01516. https://doi.org/https://doi.org/10.1128/genomeA.01544-16

Jia N, Ding MZ, Gao F, Yuan YJ (2016) Comparative genomics analysis of the companion mechanisms of *Bacillus thuringiensis* Bc601 and *Bacillus endophyticus* Hbe603 in bacterial consortium. Sci. Rep. 6:28794. https://doi.org/https://doi.org/10.1038/srep28794

Kalfon, A., Larget-Thiéry, I. Charles, JF, Barjac, H (1983) Growth, sporulation and larvicidal activity of *Bacillus sphaericus*. Appl. Microbiol. Biotechnol. 18:68-173. http://dx.doi.org/10.1007/BF00498040

Kaze, M, Brooks, L, Sistrom, M (2021) Antibiotic Resistance in *Bacillus*-based Biopesticide Products. bioRxiv, 2021-03. https://doi.org/10.1101/2021.03.15.435560

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647-1649. https://doi.org/https://doi.org/10.1093/bioinformatics/bts199

Kim, MJ, Han, JK, Park, JS, Lee, JS, Lee, SH, Cho, JI, Kim, KS (2015) Various enterotoxin and other virulence factor genes widespreadamong Bacillus cereus and *Bacillus thuringiensis* strains. J Microbiol Biotechnol 25:872-879. https://doi.org/10.4014/jmb.1502.02003

Khorramnejad A, Gomis-Cebolla J, Talaei-Hassanlouei R, Bel Y, Escriche B (2020) Genomics and proteomics analyses revealed novel candidate pesticidal proteins in a lepidopteran-toxic *Bacillus thuringiensis* strain. Toxins (Basel) 12:673. https://doi.org/https://doi.org/10.3390/toxins12110673

Konecka, E, Baranek, J, Kaznowski, A, Ziemnicka, J, Ziemnicki, K (2012) Interaction between crystalline proteins of two *Bacillus thuringiensis* strains against Spodoptera exigua. Entomol. Exp. Appl. 143:148-154. https://doi.org/10.1111/j.1570-7458.2012.01254.x

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. 35:1547. https://doi.org/https://doi.org/10.1093/molbev/msy096

Lacey LA, Grzywacz D, Shapiro-Ilan DI, Frutos R, Brownbridge M, Goettel MS (2015) Insect pathogens as biological control agents: Back to the future. J. Invertebr. Pathol. 132:1-41. https://doi.org/https://doi.org/10.1016/j.jip.2015.07.009

Lazarte, JN, Valacco, M P, Moreno, S, Salerno, GL, Berón, CM (2021) Molecular characterization of a *Bacillus thuringiensis* strain from Argentina, toxic against Lepidoptera and Coleoptera, based on its whole-genome and Cry protein analysis. J. Invertebr. Pathol., 183:107563. https://doi.org/10.1016/j.jip.2021.107563

Lechuga A, Lood C, Salas M, Van Noort V, Lavigne R, Redrejo-Rodríguez M (2020) Completed Genomic Sequence of *Bacillus thuringiensis* HER1410 reveals a Cry-containing chromosome, two megaplasmids, and an integrative plasmidial prophage. G3 Genes|Genomes|Genetics 10:2927-2939. https://doi.org/https://doi.org/10.1534/g3.120.401361

Li, X, Ding, X, Xia, L, Sun, Y, Yuan, C, Yin, J (2012) Proteomic analysis *of Bacillus thuringiensis* strain 4.0718 at different growth phases. Science World Journal. https://doi.org/10.1100/2012/798739

Liu H, Zheng J, Bo D, Yu Y, Ye W, Peng D, Sun M (2021) BtToxin_Digger: a comprehensive and high-throughput pipeline for mining toxin protein genes from *Bacillus thuringiensis*. Bioinformatics 38:250-251. https://doi.org/https://doi.org/10.1093/bioinformatics/btab506

Liu X, Zuo M, Wang T, Sun Y, Liu S, Hu S, He H, Yang Q, Rang J, Quan M, Xia L, Ding X (2015) Proteomic analysis of the influence of Cu^{2+} on the crystal protein production of *Bacillus thuringiensis* X022. Microb. Cell Factories 14, 153. https://doi.org/https://doi.org/10.1186/s12934-015-0339-9

Liu B, Zheng D, Jin Q, Chen L, Yang J (2019) VFDB 2019: a comparative pathogenomic platform with an interactive web interface. Nucleic Acids Res 47:D687–D692. https://doi.org/10.1093/nar/gky1080

Luna, VA, King, DS, Gulledge, J, Cannons, AC, Amuso, PT, Cattani, J (2007) Susceptibility of *Bacillus anthracis*, *Bacillus cereus*, *Bacillus mycoides*, *Bacillus pseudomycoides* and *Bacillus thuringiensis* to 24 antimicrobials using Sensititre® automated microbroth dilution and Etest® agar gradient diffusion methods. J. Antimicrob. Chemothe., 60: 555-567. https://doi.org/10.1093/jac/dkm213

Ma, W, Chen, H, Jiang, X, Wang, J, Gelbič, I, Guan, X, Zhang, L (2020) Whole genome sequence analysis of the mosquitocidal *Bacillus thuringiensis* LLP29. Arch Microbiol 202:1693-1700. https://doi.org/10.1007/s00203-020-01875-2

Macedo CL, Martins ÉS, Macedo LL, Santos AC, Praça LB, Góis LA, Monnerat RG (2012) Seleção e caracterização de estirpes de *Bacillus thuringiensis* eficientes contra a *Diatraea saccharalis* (Lepidoptera: Crambidae). Pesq. Agropecu. Bras. 47:1759-1765. https://doi.org/https://doi.org/10.1590/S0100-204X2012001200012

Malovichko YV, Nizhnikov AA, Antonets KS (2019) Repertoire of the *Bacillus thuringiensis* virulence factors unrelated to major classes of protein toxins and its role in specificity of host-pathogen interactions. Toxins (Basel) 11:347. https://www.mdpi.com/2072-6651/11/6/347

Mendonça EG, de Almeida Barros R, Cordeiro G, da Silva CR, Campos WG, de Oliveira JA, de Almeida Oliveira MG (2020). Larval development and proteolytic activity of *Anticarsia gemmatalis* Hübner (Lepidoptera: Noctuidae) exposed to different soybean protease inhibitors. Arch. Insect Biochem. Physiol. 103:e21637. https://doi.org/https://doi.org/10.1002/arch.21637

Miyoshi S, Shinoda S (2000) Microbial metalloproteases and pathogenesis. Microb. Infect. 2:91-98. https://doi.org/https://doi.org/10.1016/S1286-4579(00)00280-X

Monnerat RG, Batista AC, de Medeiros PT, Martins ÉS, Melatti VM, Praça LB, Dumas VF, Morinaga C, Demo C, Gomes ACM, Falcão R, Siqueira CB, Silva-Werneck JO, Berry C (2007) Screening of Brazilian *Bacillus thuringiensis* isolates active against *Spodoptera frugiperda*, *Plutella xylostella* and *Anticarsia gemmatalis*. Biol. Control 41:291-295. https://doi.org/https://doi.org/10.1016/j.biocontrol.2006.11.008

Monnerat, RG, Silva, SF, Silva-Werneck, JO (2001) Catálogo do banco de germoplasma de bactérias do gênero *Bacillus*. Brasília, DF: Embrapa Recursos Genéticos e Biotecnologia, 65 p. (Embrapa Recursos Gene´ticos e Biotecnologia. Documentos, 60)

Moscardi, F, Corrêa-Ferreira, BS, Corso, IC (2012) Pragas que atacam plântulas, hastes e pecíolos da soja. In: Hoffmann, C.B.; Corrêa-Ferreira, B.S.; Moscardi, F. (Ed.). Soja: manejo integrado de insetos e outros artrópodes-praga. Brasília, DF: Embrapa, 145-168

Mounsef, J. R., Salameh, D., kallassy Awad, M., Chamy, L., Brandam, C., Lteif, R. (2014) A simple method for the separation of *Bacillus thuringiensis* spores and crystals. J. Microbiol. Methods, 107:147-149. https://doi.org/10.1016/j.mimet.2014.10.003

Palma L, Muñoz D, Berry C, Murillo J, Caballero P (2014) *Bacillus thuringiensis* Toxins: An overview of their biocidal activity. Toxins (Basel) 6:3296-3325. https://www.mdpi.com/2072-6651/6/12/3296

Pezenti LF, Sosa-Gómez DR, de Souza RF, Vilas-Boas LA, Gonçalves KB, da Silva CR, Vilas-Bôas GT, Baranoski A, Mantovani MS, da Rosa R (2021) Transcriptional profiling analysis of susceptible and resistant strains of *Anticarsia gemmatalis* and their response to *Bacillus thuringiensis*. Genomics 113:2264-2275. https://doi.org/https://doi.org/10.1016/j.ygeno.2021.05.012

Pinheiro DH, Valicente FH (2021) Identification of *Bacillus thuringiensis* strains for the management of lepidopteran pests. Neotrop. Entomol. 50:804-811. https://doi.org/https://doi.org/10.1007/s13744-021-00896-w

Pohare MB, Wagh SG, Udayasuriyan V (2021) *Bacillus thuringiensis* as Potential Biocontrol Agent for Sustainable Agriculture, in: Yadav, A.N., Singh, J., Singh, C., Yadav, N. (Eds.), Current Trends in Microbial Biotechnology for Sustainable Agriculture. Springer Singapore, Singapore, pp. 439-468. https://doi.org/10.1007/978-981-15-6949-4_18

Praça LB, Batista AC, Martins ÉS, Siqueira CB, Dias DG, Gomes AC, Falcão R, Monnerat RG (2004) Estirpes de *Bacillus thuringiensis* efetivas contra insetos das ordens Lepidoptera, Coleoptera e Diptera. Pesq. Agropec. Bras. 39:11-16. https://doi.org/https://doi.org/10.1590/S0100-204X2004000100002

Quan M, Xie J, Liu X, Li Y, Rang J, Zhang T, Zhou F, Xia L, Hu S, Sun Y, Ding X (2016) Comparative analysis of genomics and proteomics in the new isolated *Bacillus thuringiensis* X022 revealed the metabolic regulation mechanism of carbon flux following Cu^{2+} treatment. Front. Microbiol. 7.792. https://doi.org/10.3389/fmicb.2016.00792

Rang J, He H, Wang T, Ding X, Zuo M, Quan M, Sun Y, Yu Z, Hu S, Xia L (2015) Comparative analysis of genomics and proteomics in *Bacillus thuringiensis* 4.0718. PLoS One 10:e0119065. https://doi.org/https://doi.org/10.1371/journal.pone.0119065

Richter M, Rosselló-Móra R, (2009) Shifting the genomic gold standard for the prokaryotic species definition. Proc. Natl. Acad. Sci. 106:19126-19131. https://doi.org/https://doi.org/10.1073/pnas.0906412106

Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J (2015) JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. Bioinformatics 32:929-931. https://doi.org/https://doi.org/10.1093/bioinformatics/btv681

Robertson JL, Jones MM, Olguin E, Alberts B (2017) Bioassays with arthropods, 3rd edn. CRC Press, Taylor & Francis Group, Boca Raton

Sanahuja G, Banakar R, Twyman RM, Capell T, Christou P (2011) *Bacillus thuringiensis*: a century of research, development and commercial applications. Plant Biotechnol. J. 9:283-300. https://doi.org/https://doi.org/10.1111/j.1467-7652.2011.00595.x

Santos Liu B, Zheng D, Jin Q, Chen L, Yang, J (2018) VFDB 2019: a comparative pathogenomic platform with an interactive web interface. Nucleic. Acids. Res. 47:D687-D692. https://doi.org/https://doi.org/10.1093/nar/gky1080

Santos MD, Lima DB, Fischer JS, Clasen MA, Kurt LU, Camillo-Andrade AC, Monteiro LC, de Aquino PF, Neves-Ferreira AG, Valente RH, Trugilho MR, Brunoro, GV, Souza TA, Santos RM, Batista M, Gozzo FC, Durán R, Yates JR, Barbosa VC, Carvalho PC (2022) Simple, efficient and thorough shotgun proteomic analysis with PatternLab V. Nat. Protoc. 17:1553-1578. https://doi.org/https://doi.org/10.1038/s41596-022-00690-x

Sathyan, T, Jayakanthan, M, Mohankumar, S, Balasubramani, V, Kokiladevi, E, Ravikesavan, R, Sathiah, N (2022) Genome profiling of an indigenous *Bacillus thuringiensis* isolate, T405 toxic against the fall armyworm, *Spodoptera frugiperda* (JE Smith) (Lepidoptera: Noctuidae). Microb. Pathog., 173:105820. https://doi.org/10.1016/j.micpath.2022.105820

Schmidt FG, Monnerat RG, Borges M, Carvalho R (2001) Metodologia de criação de insetos para a avaliação de agentes entomopatogênicos, in: Biotecnologia., E.R.G.e. (Ed.), Circular Técnica, Brasília, DF

Schnepf E, Crickmore N, Rie JV, Lereclus D, Baum J, Feitelson J, Zeigler DR, Dean DH (1998) *Bacillus thuringiensis* and its pesticidal crystal proteins. Microbiol. Mol. Biol Rev. 62:775-806. https://doi.org/https://doi.org/10.1128/MMBR.62.3.775-806.1998

Sena da Silva IH, Mueller de Freitas M, Polanczyk RA (2022) *Bacillus thuringiensis*, a remarkable biopesticide: from lab to the field, in: Rakshit, A., Meena, V.S., Abhilash, P.C., Sarma, B.K., Singh, H.B., Fraceto, L., Parihar, M., Singh, A.K. (Eds.), Biopesticides. Woodhead Publishing, pp. 117-131. https://doi.org/https://doi.org/10.1016/B978-0-12-823355-9.00021-3

Silva, MTB (1995) Associação de Baculovirus anticarsia com subdosagem de inseticidas no controle de *Anticarsia gemmatalis* (Hübner, 1818). Ciência Rural, 25: 353-358.

Singh D, Samiksha Thayil SM, Sohal SK, Kesavan AK (2021) Exploration of insecticidal potential of Cry protein purified from *Bacillus thuringiensis* VIID1. Int. J. Biol. Macromol. 174:362-369. https://doi.org/https://doi.org/10.1016/j.ijbiomac.2021.01.143

Sreshty, MAL, Kumar, KP, Murty, USN (2011) Synergism between wild-type *Bacillus thuringiensis* subsp. *israelensis* and *B. sphaericus* strains: A study based on isobolographic analysis and histopathology. Act. Tropica. 118:14-20. https://doi.org/10.1016/j.actatropica.2010.12.012

Srikanth J, Subramonian N, Premachandran MN (2011) Advances in Transgenic Research for Insect Resistance in Sugarcane. Trop. Plant Biol. 4:52-61. https://doi.org/https://doi.org/10.1007/s12042-011-9077-2

Sun D, Zhu L, Guo L, Wang S, Wu Q, Crickmore N, Zhou X, Bravo A, Soberón M, Guo Z, Zhang Y (2022) A versatile contribution of both aminopeptidases N and ABC transporters to Bt Cry1Ac toxicity in the diamondback moth. BMC Biol. 20:33. https://doi.org/https://doi.org/10.1186/s12915-022-01226-1

Tailor R, Tippett J, Gibb G, Pells S, Jordan L, Ely S (1992) Identification and characterization of a novel *Bacillus thuringiensis* δ-endotoxin entomocidal to coleopteran and lepidopteran larvae. Mol. Microbiol. 6:1211-1217. https://doi.org/https://doi.org/10.1111/j.1365-2958.1992.tb01560.x

Tan SY, Cayabyab BF, Alcantara EP, Ibrahim YB, Huang F, Blankenship EE, Siegfried BD (2011) Comparative susceptibility of *Ostrinia furnacalis*, *Ostrinia nubilalis* and *Diatraea saccharalis* (Lepidoptera: Crambidae) to *Bacillus thuringiensis* Cry1 toxins. Crop Protect. 30:1184-1189. https://doi.org/https://doi.org/10.1016/j.cropro.2011.05.009

Wu D, He J, Gong Y, Chen D, Zhu X, Qiu N, Sun M, Li M, Yu Z (2011) Proteomic analysis reveals the strategies of *Bacillus thuringiensis* YBT-1520 for survival under long-term heat stress. Proteomics 11:2580-2591. https://doi.org/https://doi.org/10.1002/pmic.201000392

Xie J, Peng J, Yi Z, Zhao X, Li S, Zhang T, Quan M, Yang S, Lu J, Zhou P, Xia L, Ding X (2019) Role of hsp20 in the production of spores and insecticidal crystal proteins in *Bacillus thuringiensis*. Front. Microbiol. 10:2059 https://doi.org/https://doi.org/10.3389/fmicb.2019.0205

Ye J, Zhang Y, Cui H, Liu J, Wu Y, Cheng Y, Xu H, Huang X, Li S, Zhou A, Zhang X, Bolund L, Chen Q, Wang J, Yang H, Fang L, Shi C (2018) WEGO 2.0: a web tool for analyzing and plotting GO annotations, 2018 update. Nucleic Acids Res. 46:W71-W75. https://doi.org/https://doi.org/10.1093/nar/gky400

Yin, J, Ding, X, Xia, L, Yu, Z, Lv, Y, Hu, S, Xiao, X (2011) Transcription of gene in an acrystalliferous strain of *Bacillus thuringiensis* XBU001 positively regulated by the metalloprotease camelysin gene at the onset of stationary phase. FEMS Microbiol. Lett., 318: 92-100. https://doi.org/10.1111/j.1574-6968.2011.02247.x

Zghal, RZ, Ghedira, K, Elleuch, J, Kharrat, M, Tounsi, S (2018) Genome sequence analysis of a novel *Bacillus thuringiensis* strain BLB406 active against *Aedes aegypti* larvae, a novel potential bioinsecticide. J. Biol. Macromol., 116:1153-1162. https://doi.org/10.1016/j.ijbiomac.2018.05.119

Zhou H, Zhang J, Shao Y, Wang J, Xu W, Liu Y, Yu S, Ye Q, Pang R, Wu S, Gu Q, Xue L, Zhang J, Li H, Wu Q, Ding, Y (2022) Development of a high resolution melting method based on a novel molecular target for discrimination between *Bacillus cereus* and *Bacillus thuringiensis*. Food. Res. Int. 151, 110845. https://doi.org/https://doi.org/10.1016/j.foodres.2021.110845

Zhu, L, Peng, D, Wang, Y et al (2015) Genomic and transcriptomic insights into the efficient entomopathogenicity of *Bacillus thuringiensis*. Sci Rep. 5:14129. https://doi.org/10.1038/srep14129

4 CONCLUSIONS

This work demonstrated that *Bt* strains isolated from the Tocantins are alternatives for dipteran and lepidopteran insects control. Between Bt strains with the most toxicity against Aedes aegypti and Culex quinquefasciatus, *Bt* T0131 and *Bt* TOD651 showed high larvicidal activity. *Bt* UFT038 showed high toxicity against Spodoptera cosmioides. *Bt* TOL651 may be used to control Diatreae saccharalis and Anticarsia germmatalis and was more effective than commercial strain.

The genomic analysis detected all *cry*, *cyt*, *mpp*, and *vip* genes content that may be the main responsible for the pathogenicity of *Bt* strains. However, no novel *cry* protein was detected in the genomes.

The genomic-proteomic analysis of two bacteria revealed the lack of expression of genes considered cryptics. Besides, virulence factors InhA1 and Mppe were detected in the spore-crystal mixtures and may contribute to the pathogenicity of *Bt* strains.

The study of the insecticide potential of individual proteins will be necessary for better understanding the role of each one in the *Bt* strains' pathogenicity. Besides, it is needed will elucidate the role-play of the Mppe enzyme in the Cry and Cyt mixture.

Finally, the genomic and proteomic results showed to be an effective tool for the applicating in the study of the bioinsecticidal potential of *Bt* strains.

5 APPENDICES

5.1 Appendice A- Supplemental material of Chapter 2

Table S1. Virulence factors identified in the chromosome draft sequence of *Bt* TOD651 from the VFDB database.

VFclass	Virulence factors	Related genes	TOD651 (Prediction)	Start-End	Length	Strand
Enzyme	Immune inhibitor A metalloproteinase	*Undetermined*	orf00902	825173-827521	2349	Plus
		inhA	orf00231	234646-237045	2400	Plus
			orf02721	2432400-2434784	2385	Minus
	Phosphatidylcholine-preferring phospholipase C (PC-PLC)	*plcA*	orf00237	241223-242074	852	Plus
	Phosphatidylinositol-specific phospholipase C (PI-PLC)	*piplc*	orf03644	3229668-3230657	990	Minus
	Sphingomyelinase (SMase)	*sph*	orf00238	242151-243152	1002	Plus
Immune evasion	Polysaccharide capsule (Pc)	*Undetermined*	orf05247	4674211- 4675029	819	Minus
			orf05248	4675128- 4676039	912	Minus
			orf05251	4677351- 4678673	1323	Minus
			orf05259	4684071-4684949	879	Minus
			orf05260	4685189- 4685965	777	Minus
			orf05261	4686068-4686769	702	Minus
			orf05262	4686759-4687514	756	Minus
			orf05263	4687766- 4688506	741	
Iron acquisition	Bacillibactin	*dhbA*	orf01972	1752641-1753411	771	Plus
		dhbB	orf01975	1756291- 1757184	894	Plus
		dhbC	orf01973	1753438-1754637	1200	Plus
		dhbE	orf01974	1754650-1756266	1617	Plus
		dhbF	orf01976	1757218-1764372	7155	Plus
	IlsA	*ilsA*	orf00102;	100504-103575	3072	Plus
			orf00950	871952-874612	2661	Plus
	Petrobactin	*asbA*	orf01592	1427219-1429027	1809	Plus
		asbB	orf01593	1429088-1430926	1839	Plus
		asbC	orf01594	1430913-1432151	1239	Plus
		asbD	orf01595	1432187-1432423	237	Plus
		asbE	orf01596	1432474-1433430	957	Plus
		asbF	orf01597	1433468-1434361	894	Plus
Regulation	PagR-XO2	*pagR-XO2*	orf02128	1923192-1923485	294	Plus
	PlcR-PapR quorum sensing	*papR*	orf05361	4774672-4776879	2208	Plus
			orf05362	4776896-4777042	147	Minus
		plcR	orf05363	4777130-4777987	858	Minus
Toxin	Hemolysin I	*hlyI*	orf05099	4530613-4532151	1539	Minus
	Hemolysin II	*hlyII*	orf03324	2945393-2945959	567	Minus
	Hemolysin III homolog	*Undetermined*	orf05471	4891786-4892448	663	Minus
	Hemolysin III	*hlyIII*	orf01851	1650946-1651662	717	Minus
	Hemolytic enterotoxin HBL	*hblA*	orf02860	2540164-2541564	1401	Minus
			orf02862	2541690- 2543804	2115	Minus
	Non-hemolytic enterotoxin (Nhe)	*nheA*	orf01475	1329172-1330332	1161	Plus
		nheB	orf01476	1330364- 1331572	1209	Plus
		nheC	orf01477	1331671-1332759	1089	Plus
Magnesium uptake	Mg^{2+} transport	*mgtB*	orf04108	3648738-3651464	2727	Plus

Table S2.1 Peptides identified to Cry4Ba5 protein classified as unique (true) or non-unique (false).

CDS id	Peptide sequence	unique
peg.5812	NVLQNGDFESATLGWTTSDNITIQEDDPIFK	False
peg.5812	NVLQNGDFESATLGWTTSDNITIQEDDPIFK	False
peg.5812	NVLQNGDFESATLGWTTSDNITIQEDDPIFK	False
peg.5812	NVLQNGDFESATLGWTTSDNITIQEDDPIFK	False
peg.5812	NVLQNGDFESATLGWTTSDNITIQEDDPIFK	False
peg.5812	YAANSPIVLNVSYVLQGVSR	True
peg.5812	YAANSPIVLNVSYVLQGVSR	True
peg.5812	IEIIPITQSVLDETENQNLESER	True
peg.5812	IEIIPITQSVLDETENQNLESER	True
peg.5812	IEIIPITQSVLDETENQNLESER	True
peg.5812	IEIIPITQSVLDETENQNLESER	True
peg.5812	IEIIPITQSVLDETENQNLESER	True
peg.5812	KEGPGNGYVTLMDCEENQEK	True
peg.5812	DALNIGTTDYDIDQAANLVECISEELYPK	True
peg.5812	DALNIGTTDYDIDQAANLVECISEELYPK	True
peg.5812	DALNIGTTDYDIDQAANLVECISEELYPK	True
peg.5812	DALNIGTTDYDIDQAANLVECISEELYPK	True
peg.5812	DALNIGTTDYDIDQAANLVECISEELYPK	True
peg.5812	DALNIGTTDYDIDQAANLVECISEELYPK	True
peg.5812	DALNIGTTDYDIDQAANLVECISEELYPK	True
peg.5812	DALNIGTTDYDIDQAANLVECISEELYPK	True
peg.5812	DALNIGTTDYDIDQAANLVECISEELYPK	True
peg.5812	DALNIGTTDYDIDQAANLVECISEELYPK	True
peg.5812	YGEEIDAIMNVPADLNYLYPSTFDCEGSNR	True
peg.5812	YGEEIDAIMNVPADLNYLYPSTFDCEGSNR	True
peg.5812	YGEEIDAIMNVPADLNYLYPSTFDCEGSNR	True
peg.5812	YGEEIDAIMNVPADLNYLYPSTFDCEGSNR	True
peg.5812	YGEEIDAIM[15.9949]NVPADLNYLYPSTFDCEGSNR	True
peg.5812	YGEEIDAIM[15.9949]NVPADLNYLYPSTFDCEGSNR	True
peg.5812	DWLAMCENNQQYGVNPAAINSSSVSTALK	True
peg.5812	FVNPPAGTVLTVLSAVLPILWPTNTPTPER	True
peg.5812	FVNPPAGTVLTVLSAVLPILWPTNTPTPER	True
peg.5812	TDVIDYNSNR	True
peg.5812	TDVIDYNSNR	True
peg.5812	TDVIDYNSNR	True
peg.5812	TDVIDYNSNR	True
peg.5812	TDVIDYNSNR	True
peg.5812	WNDQM[15.9949]EAK	True
peg.5812	WNDQM[15.9949]EAK	True
peg.5812	WNDQM[15.9949]EAK	True
peg.5812	WNDQM[15.9949]EAK	True

peg.5812	WNDQM[15.9949]EAK	True
peg.5812	WNDQM[15.9949]EAK	True
peg.5812	WNDQM[15.9949]EAK	True
peg.5812	WNDQMEAK	True
peg.5812	WNDQMEAK	True
peg.5812	EIYTALVESPSSK	True
peg.5812	EIYTALVESPSSK	True
peg.5812	EIYTALVESPSSK	True
peg.5812	DYLDQYTTK	True
peg.5812	DYLDQYTTK	True
peg.5812	TSITDTSSPSNR	True
peg.5812	TSITDTSSPSNR	True
peg.5812	TSITDTSSPSNR	True
peg.5812	SNFLNATAK	True
peg.5812	SNFLNATAK	True
peg.5812	SIAALEAALTR	True
peg.5812	SIAALEAALTR	True
peg.5812	SIAALEAALTR	True
peg.5812	SIAALEAALTR	True
peg.5812	SIAALEAALTR	True
peg.5812	DVELVVSR	True
peg.5812	DVELVVSR	True
peg.5812	DVELVVSR	True
peg.5812	DVELVVSR	True
peg.5812	TVDVFPDTDR	True
peg.5812	TVDVFPDTDR	True
peg.5812	TVDVFPDTDR	True
peg.5812	TVDVFPDTDR	True
peg.5812	TVDVFPDTDR	True
peg.5812	TVDVFPDTDR	True
peg.5812	TSIFNDPTR	True
peg.5812	TSIFNDPTR	True
peg.5812	TSIFNDPTR	True
peg.5812	SNGQWITFNDYK	True
peg.5812	SNGQWITFNDYK	True
peg.5812	SNGQWITFNDYK	True
peg.5812	GPGHTGGDLVALTSN[0.9840]GTLSGR	True
peg.5812	GPGHTGGDLVALTSNGTLSGR	True
peg.5812	CETSAVPANIGNTSDMLYSCQYDTGK	True
peg.5812	CETSAVPANIGNTSDMLYSCQYDTGK	True
peg.5812	SAGDQLYNTMVQYTK	True
peg.5812	SAGDQLYNTM[15.9949]VQYTK	True
peg.5812	EGPGNGYVTLMDCEENQEK	True
peg.5812	EGPGNGYVTLMDCEENQEK	True
peg.5812	RVDFWTNTIYQDLR	True

peg.5812	RVDFWTNTIYQDLR	True
peg.5812	TAVITQFNLTSAK	True
peg.5812	IDGTLASYNSNITPTPEGLR	True
peg.5812	IDGTLASYNSNITPTPEGLR	True
peg.5812	IDGTLASYNSNITPTPEGLR	True
peg.5812	SNGQWITFNDYKR	True
peg.5812	SNGQWITFNDYKR	True
peg.5812	VDFWTNTIYQDLR	True
peg.5812	VDFWTNTIYQDLR	True
peg.5812	DIDGTIFPTYIFQK	True
peg.5812	DIDGTIFPTYIFQK	True
peg.5812	DIDGTIFPTYIFQK	True
peg.5812	DIDGTIFPTYIFQK	True
peg.5812	DIDGTIFPTYIFQK	True
peg.5812	DIDGTIFPTYIFQK	True
peg.5812	DIDGTIFPTYIFQK	True
peg.5812	DIDGTIFPTYIFQK	True
peg.5812	DIDGTIFPTYIFQ[0.9840]K	True
peg.5812	DGLINAQEWSLAR	True
peg.5812	DGLINAQEWSLAR	True
peg.5812	SETQQAYDVAK	True
peg.5812	LTFTSCEEGYITK	True
peg.5812	LTFTSCEEGYITK	True
peg.5812	KWNDQM[15.9949]EAK	True
peg.5812	KWNDQM[15.9949]EAK	True
peg.5812	KWNDQMEAK	True
peg.5812	YKDPFDAIVPMR	True
peg.5812	YKDPFDAIVPMR	True
peg.5812	YKDPFDAIVPM[15.9949]R	True
peg.5812	YKDPFDAIVPM[15.9949]R	True
peg.5812	YKDPFDAIVPM[15.9949]R	True
peg.5812	RYPADKIDNTK	True
peg.5812	RSETQQAYDVAK	True
peg.5812	RSETQQAYDVAK	True
peg.5812	RSETQQAYDVAK	True
peg.5812	RSETQQAYDVAK	True
peg.5812	RSETQQAYDVAK	True
peg.5812	M[15.9949]NSGYPLANDLQGSMK	True
peg.5812	DVHLFTWLK	True
peg.5812	SNFLN[0.9840]ATAK	True
peg.5812	EPNNQSYR	True
peg.5812	TVDVFPDTDRVR	True
peg.5812	TVDVFPDTDRVR	True
peg.5812	GTTISTESTFSRPNNIIPTDLK	True
peg.5812	DPFDAIVPMR	True
peg.5812	YPADKIDNTK	True

peg.5812	ISSPDGYASLDNLEVIEEGPIDGEALSR	True
peg.5812	CETSAVPANIGNTSDMLYSCQYDTGKK	True
peg.5812	CETSAVPANIGNTSDMLYSCQYDTGKK	True
peg.5812	CETSAVPANIGNTSDM[15.9949]LYSCQYDTGKK	True
peg.5812	CETSAVPANIGNTSDM[15.9949]LYSCQYDTGKK	True
peg.5812	CETSAVPANIGNTSDM[15.9949]LYSCQYDTGKK	True
peg.5812	CETSAVPANIGNTSDM[15.9949]LYSCQYDTGK	True
peg.5812	EYIAHSITWYNK	True
peg.5812	TTFFGFSTNENTPNQPTVNDYTHILSYIK	True
peg.5812	GHYLHM[15.9949]SGAR	True
peg.5812	MLLLDEVK	False
peg.5812	RVDFWTNTIYQDLR	True
peg.5812	RYPADKIDNTK	True
peg.5812	GHYLHM[15.9949]SGAR	True
peg.5812	REPNNQSYR	True
peg.5812	ISSPDGYASLDNLEVIEEGPIDGEALSR	True
peg.5812	YPADKIDNTK	True
peg.5812	GPGHTGGDLVALTSN[0.9840]GTLSGR	True
peg.5812	TVDVFPDTDRVR	True
peg.5812	ISSPDGYASLDNLEVIEEGPIDGEALSR	True
peg.5812	VDFWTNTIYQDLR	True
peg.5812	TVDVFPDTDRVR	True
peg.5812	GPGHTGGDLVALTSN[0.9840]GTLSGR	True
peg.5812	NTNYKDWLAMCENNQQYGVNPAAINSSSVSTALK	True
peg.5812	RVDFWTNTIYQDLR	True
peg.5812	YKDPFDAIVPMR	True
peg.5812	TVDVFPDTDRVR	True
peg.5812	ISSPDGYASLDN[0.9840]LEVIEEGPIDGEALSR	True
peg.5812	RYPADKIDNTK	True
peg.5812	DVHLFTWLK	True
peg.5812	YKDPFDAIVPM[15.9949]R	True
peg.5812	YKDPFDAIVPMR	True
peg.5812	TVDVFPDTDRVR	True
peg.5812	SAGDQLYNTMVQYTK	True
peg.5812	YKDPFDAIVPM[15.9949]R	True
peg.5812	YKDPFDAIVPMR	True
peg.5812	RYPADKIDNTK	True
peg.5812	YKDPFDAIVPMR	True
peg.5812	GPGHTGGDLVALTSNGTLSGR	True

Table S2.2 Peptides identified to Cry4Aa4 protein classified as unique (true) or non-unique (false).

CDS id	Peptide sequence	unique
peg.1190	SSVFGNHNVTDKLK	True
peg.1190	NEYEIFNAPSN[0.9840]GFSK	False
peg.1190	YLNDYNNISK	True
peg.1190	YLNDYNNISK	True
peg.1190	YLNDYNNISK	True
peg.1190	YLNDYNNISK	True
peg.1190	FAPNQNISLVFNR	True
peg.1190	FAPNQNISLVFNR	True
peg.1190	FAPNQNISLVFNR	True
peg.1190	RPHLFTWLDSLNFYEK	True
peg.1190	RPHLFTWLDSLNFYEK	True
peg.1190	RPHLFTWLDSLNFYEK	True
peg.1190	RPHLFTWLDSLNFYEK	True
peg.1190	YASNGSANTR	True
peg.1190	YASN[0.9840]GSANTR	True
peg.1190	ENQGNPTLFPTYDNYSHILSFIK	True
peg.1190	ENQGNPTLFPTYDNYSHILSFIK	True
peg.1190	ENQGNPTLFPTYDNYSHILSFIK	True
peg.1190	ENQGNPTLFPTYDNYSHILSFIK	True
peg.1190	ENQGNPTLFPTYDNYSHILSFIK	True
peg.1190	ENQGNPTLFPTYDNYSHILSFIK	True
peg.1190	NIFGLPILK	True
peg.1190	NIFGLPILK	True
peg.1190	NIFGLPILK	True
peg.1190	NIFGLPILK	True
peg.1190	NIFGLPILK	True
peg.1190	NIFGLPILK	True
peg.1190	NIFGLPILK	True
peg.1190	NIFGLPILK	True
peg.1190	YPIGVQSELTR	True
peg.1190	YPIGVQSELTR	True
peg.1190	YPIGVQSELTR	True
peg.1190	GHYLHM[15.9949]	True
peg.1190	SDVYTNTTVLIDK	True
peg.1190	SDVYTNTTVLIDK	True
peg.1190	SDVYTNTTVLIDK	True
peg.1190	EIASTYISNANK	True
peg.1190	EIASTYISNANK	True
peg.1190	SSVFGNHNVTDK	True

peg.1190	SSVFGNHNVTDK	True
peg.1190	SSVFGNHNVTDK	True
peg.1190	SLSIPATYK	True
peg.1190	IEFLPITR	True
peg.1190	IEFLPITR	True
peg.1190	ELTAGSGQITYDVNK	True
peg.1190	ELTAGSGQITYDVNK	True
peg.1190	ELTAGSGQITYDVNK	True
peg.1190	ELTAGSGQITYDVNK	True
peg.1190	ELTAGSGQITYDVNK	True
peg.1190	ELTAGSGQITYDVNK	True
peg.1190	SDVYTNTTVLIDKIEFLPITR	True
peg.1190	SDVYTNTTVLIDKIEFLPITR	True
peg.1190	SDVYTNTTVLIDKIEFLPITR	True
peg.1190	SDVYTNTTVLIDKIEFLPITR	True
peg.1190	SDVYTNTTVLIDKIEFLPITR	True
peg.1190	SDVYTNTTVLIDKIEFLPITR	True
peg.1190	SDVYTNTTVLIDKIEFLPITR	True
peg.1190	SDVYTNTTVLIDKIEFLPITR	True
peg.1190	TTPDSNLDGNINWNTYNTYR	True
peg.1190	TTPDSNLDGNINWNTYNTYR	True
peg.1190	TTPDSNLDGNINWNTYNTYR	True
peg.1190	TTPDSNLDGNINWNTYNTYR	True
peg.1190	YYDFQYQEDSLTR	True
peg.1190	YYDFQYQEDSLTR	True
peg.1190	YYDFQYQEDSLTR	True
peg.1190	YYDFQYQEDSLTR	True
peg.1190	YYDFQYQEDSLTR	True
peg.1190	YYDFQYQEDSLTR	True
peg.1190	QFDYLEPLPTAIDYYPVLTK	True
peg.1190	QFDYLEPLPTAIDYYPVLTK	True
peg.1190	QFDYLEPLPTAIDYYPVLTK	True
peg.1190	QFDYLEPLPTAIDYYPVLTK	True
peg.1190	QFDYLEPLPTAIDYYPVLTK	True
peg.1190	QFDYLEPLPTAIDYYPVLTK	True
peg.1190	QFDYLEPLPTAIDYYPVLTK	True
peg.1190	QFDYLEPLPTAIDYYPVLTK	True
peg.1190	QFDYLEPLPTAIDYYPVLTK	True
peg.1190	QFDYLEPLPTAIDYYPVLTK	True
peg.1190	QFDYLEPLPTAIDYYPVLTK	True
peg.1190	QFDYLEPLPTAIDYYPVLTK	True
peg.1190	QFDYLEPLPTAIDYYPVLTK	True
peg.1190	QFDYLEPLPTAIDYYPVLTK	True

peg.1190	QFDYLEPLPTAIDYYPVLTK	True
peg.1190	QFDYLEPLPTAIDYYPVLTK	True
peg.1190	QFDYLEPLPTAIDYYPVLTK	True
peg.1190	QFDYLEPLPTAIDYYPVLTK	True
peg.1190	QFDYLEPLPTAIDYYPVLTK	True
peg.1190	QFDYLEPLPTAIDYYPVLTK	True
peg.1190	QFDYLEPLPTAIDYYPVLTK	True
peg.1190	QFDYLEPLPTAIDYYPVLTK	True
peg.1190	QFDYLEPLPTAIDYYPVLTK	True
peg.1190	AIEDYTNYCVTTYK	True
peg.1190	AIEDYTNYCVTTYK	True
peg.1190	AIEDYTNYCVTTYK	True
peg.1190	AIEDYTNYCVTTYK	True
peg.1190	AIEDYTN[0.9840]YCVTTYK	True
peg.1190	AIEDYTNYCVTTYKK	True
peg.1190	AIEDYTNYCVTTYKK	True
peg.1190	MDFFITN[0.9840]GTR	True
peg.1190	MDFFITN[0.9840]GTR	True
peg.1190	MDFFITNGTR	True
peg.1190	SFNVISTYHNHLK	True
peg.1190	SFNVISTYHNHLK	True
peg.1190	SFNVISTYHNHLK	True
peg.1190	SFNVISTYHNHLK	True
peg.1190	VVQGPGHTGGDLIDFK	True
peg.1190	NTIYTHLTTQIPAVK	True
peg.1190	NTIYTHLTTQIPAVK	True
peg.1190	TWENNPNPQNTQDVR	True
peg.1190	TWENNPNPQNTQDVR	True
peg.1190	TWENNPNPQNTQDVR	True
peg.1190	TWENNPNPQNTQDVR	True
peg.1190	TWENNPNPQNTQDVR	True
peg.1190	YKDFQYLEFSNEVK	True
peg.1190	M[15.9949]TTAVLDLVALFPNYDVGK	True
peg.1190	VVQGPGHTGGDLIDFKDHFK	True
peg.1190	VVQGPGHTGGDLIDFKDHFK	True
peg.1190	YPLANKPNQPLK	False
peg.1190	YPLANKPNQPLK	False
peg.1190	YPLANKPNQPLK	False
peg.1190	DFQYLEFSNEVK	True
peg.1190	EIYQVLNFEESPYK	True
peg.1190	SLGLATNIYIFLLNVISLDNK	True
peg.1190	SLGLATNIYIFLLNVISLDNK	True
peg.1190	NTLQSELTDYDIDQAANLVECISEELYPK	True

peg.1190	NTLQSELTDYDIDQAANLVECISEELYPK	True
peg.1190	NTLQSELTDYDIDQAANLVECISEELYPK	True
peg.1190	NTLQSELTDYDIDQAANLVECISEELYPK	True
peg.1190	NTLQSELTDYDIDQAANLVECISEELYPK	True
peg.1190	NTLQSELTDYDIDQAANLVECISEELYPK	True
peg.1190	NTLQSELTDYDIDQAANLVECISEELYPK	True
peg.1190	NTLQSELTDYDIDQAANLVECISEELYPK	True
peg.1190	NTLQSELTDYDIDQAANLVECISEELYPK	True
peg.1190	Q[0.9840]FDYLEPLPTAIDYYPVLTK	True
peg.1190	Q[0.9840]FDYLEPLPTAIDYYPVLTK	True
peg.1190	Q[0.9840]FDYLEPLPTAIDYYPVLTK	True
peg.1190	AVINLSIPGVAELGMALNPTFSGTDYTNLK	True
peg.1190	AVINLSIPGVAELGMALNPTFSGTDYTNLK	True
peg.1190	AVINLSIPGVAELGM[15.9949]ALNPTFSGTDYTNLK	True
peg.1190	AVINLSIPGVAELGM[15.9949]ALNPTFSGTDYTNLK	True
peg.1190	SDVYTN[0.9840]TTVLIDKIEFLPITR	True
peg.1190	SDVYTN[0.9840]TTVLIDKIEFLPITR	True
peg.1190	ELTAGSGQITYDVNKNIFGLPILK	True
peg.1190	M[15.9949]TTAVLDLVALFPNYDVGKYPIGVQSELTR	True
peg.1190	M[15.9949]TTAVLDLVALFPNYDVGKYPIGVQSELTR	True
peg.1190	AQTTPNNFFTSHYNMFHYTLDNISQK	True
peg.1190	AQTTPNNFFTSHYNMFHYTLDNISQK	True
peg.1190	NVLQNGDFESATLGWTTSDNITIQEDDPIFK	False
peg.1190	NVLQNGDFESATLGWTTSDNITIQEDDPIFK	False
peg.1190	NVLQNGDFESATLGWTTSDNITIQEDDPIFK	False
peg.1190	NVLQNGDFESATLGWTTSDNITIQEDDPIFK	False
peg.1190	NVLQNGDFESATLGWTTSDNITIQEDDPIFK	False
peg.1190	SLSIPATYK	True
peg.1190	SLSIPATYK	True
peg.1190	MLLLDEVK	False
peg.1190	FEAYLK	True
peg.1190	YKDFQYLEFSNEVK	True
peg.1190	VVQGPGHTGGDLIDFKDHFK	True
peg.1190	RENQGNPTLFPTYDNYSHILSFIK	True
peg.1190	M[15.9949]TTAVLDLVALFPNYDVGK	True
peg.1190	SSVFGNHNVTDKLK	True
peg.1190	RENQGNPTLFPTYDNYSHILSFIK	True
peg.1190	VVQGPGHTGGDLIDFK	True
peg.1190	NTIYTHLTTQIPAVK	True
peg.1190	YKDFQYLEFSNEVK	True
peg.1190	QKLETVQQIINTFYANPIK	True
peg.1190	SSVFGNHNVTDK	True
peg.1190	SSVFGNHNVTDK	True

peg.1190	SSVFGNHNVTDK	True
peg.1190	VVQGPGHTGGDLIDFK	True
peg.1190	SSVFGNHNVTDKLK	True
peg.1190	TWENNPNPQNTQDVR	True
peg.1190	VVQGPGHTGGDLIDFKDHFK	True
peg.1190	TWENNPNPQNTQDVR	True
peg.1190	FAPNQNISLVFNRSDVYTNTTVLIDKIEFLPITR	True
peg.1190	VVQGPGHTGGDLIDFKDHFK	True

Table S2.3 Peptides identified to Cry11Aa3 protein classified as unique (true) or non-unique (false).

CDS id	Peptide sequence	unique
peg.1401	AISACPR	True
peg.1401	TFISNEAK	True
peg.1401	TFISNEAK	True
peg.1401	TFISNEAK	True
peg.1401	VNILNAEYR	True
peg.1401	VNILNAEYR	True
peg.1401	VNILNAEYR	True
peg.1401	VNILN[0.9840]AEYR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGNGR	True
peg.1401	GVHPIYDPSSGLTGWIGN[0.9840]GR	True
peg.1401	GVHPIYDPSSGLTGWIGN[0.9840]GR	True
peg.1401	GVHPIYDPSSGLTGWIGN[0.9840]GR	True
peg.1401	GVHPIYDPSSGLTGWIGN[0.9840]GR	True

peg.1401	GVHPIYDPSSGLTGWIGN[0.9840]GR	True
peg.1401	GVHPIYDPSSGLTGWIGN[0.9840]GR	True
peg.1401	GVHPIYDPSSGLTGWIGN[0.9840]GR	True
peg.1401	GVHPIYDPSSGLTGWIGN[0.9840]GR	True
peg.1401	GVHPIYDPSSGLTGWIGN[0.9840]GR	True
peg.1401	GVHPIYDPSSGLTGWIGN[0.9840]GR	True
peg.1401	GVHPIYDPSSGLTGWIGN[0.9840]GR	True
peg.1401	GVHPIYDPSSGLTGWIGN[0.9840]GR	True
peg.1401	GVHPIYDPSSGLTGWIGN[0.9840]GR	True
peg.1401	GVHPIYDPSSGLTGWIGN[0.9840]GR	True
peg.1401	GVHPIYDPSSGLTGWIGN[0.9840]GR	True
peg.1401	GYFLNLSGAIIQR	True
peg.1401	GYFLNLSGAIIQR	True
peg.1401	GYFLNLSGAIIQR	True
peg.1401	GYFLNLSGAIIQR	True
peg.1401	GYFLNLSGAIIQR	True
peg.1401	GYFLNLSGAIIQR	True
peg.1401	ESAFTTQINPLLK	True
peg.1401	ESAFTTQINPLLK	True
peg.1401	ESAFTTQINPLLK	True
peg.1401	ESAFTTQINPLLK	True
peg.1401	ESAFTTQINPLLK	True
peg.1401	ESAFTTQINPLLK	True
peg.1401	ESAFTTQINPLLK	True
peg.1401	ESAFTTQINPLLK	True
peg.1401	QPGFTPATAK	True
peg.1401	QPGFTPATAK	True
peg.1401	QPGFTPATAK	True
peg.1401	QPGFTPATAK	True
peg.1401	QPGFTPATAK	True
peg.1401	TEVETLINQK	True
peg.1401	TEVETLINQK	True
peg.1401	TEVETLINQK	True
peg.1401	TEVETLINQK	True
peg.1401	TEVETLINQK	True
peg.1401	TEVETLINQK	True
peg.1401	TEVETLINQK	True
peg.1401	TEVETLINQK	True
peg.1401	SHYLSETNDSYVIPALQFAEVSDR	True
peg.1401	SHYLSETNDSYVIPALQFAEVSDR	True
peg.1401	SHYLSETNDSYVIPALQFAEVSDR	True
peg.1401	SHYLSETNDSYVIPALQFAEVSDR	True
peg.1401	SHYLSETNDSYVIPALQFAEVSDR	True

peg.1401	SHYLSETNDSYVIPALQFAEVSDR	True
peg.1401	SHYLSETNDSYVIPALQFAEVSDR	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	DGILAGSAWGFTQADVDSFIK	True
peg.1401	FNYSFTNEPADIPAR	True
peg.1401	FNYSFTNEPADIPAR	True

peg.1401	FNYSFTNEPADIPAR	True
peg.1401	FNYSFTNEPADIPAR	True
peg.1401	FNYSFTNEPADIPAR	True
peg.1401	FNYSFTNEPADIPAR	True
peg.1401	FNYSFTNEPADIPAR	True
peg.1401	FNYSFTNEPADIPAR	True
peg.1401	FNYSFTNEPADIPAR	True
peg.1401	FNYSFTNEPADIPAR	True
peg.1401	FNYSFTNEPADIPAR	True
peg.1401	FNYSFTNEPADIPAR	True
peg.1401	FNYSFTNEPADIPAR	True
peg.1401	FNYSFTNEPADIPAR	True
peg.1401	FNYSFTNEPADIPAR	True
peg.1401	FNYSFTNEPADIPAR	True
peg.1401	GVSLAYNHDLTTLTYNR	True
peg.1401	GVSLAYNHDLTTLTYNR	True
peg.1401	GVSLAYNHDLTTLTYNR	True
peg.1401	GVSLAYNHDLTTLTYNR	True
peg.1401	GVSLAYNHDLTTLTYNR	True
peg.1401	GVSLAYNHDLTTLTYNR	True
peg.1401	GVSLAYNHDLTTLTYNR	True
peg.1401	GVSLAYNHDLTTLTYNR	True
peg.1401	GVSLAYNHDLTTLTYNR	True
peg.1401	GVSLAYNHDLTTLTYNR	True
peg.1401	GVSLAYNHDLTTLTYNR	True
peg.1401	GVSLAYNHDLTTLTYNR	True
peg.1401	GVSLAYNHDLTTLTYNR	True
peg.1401	GVSLAYNHDLTTLTYNR	True
peg.1401	GVSLAYNHDLTTLTYNR	True
peg.1401	GVSLAYNHDLTTLTYNR	True
peg.1401	GVSLAYNHDLTTLTYNR	True
peg.1401	KSHYLSETNDSYVIPALQFAEVSDR	True
peg.1401	KSHYLSETNDSYVIPALQFAEVSDR	True
peg.1401	KSHYLSETNDSYVIPALQFAEVSDR	True
peg.1401	KSHYLSETNDSYVIPALQFAEVSDR	True
peg.1401	KSHYLSETNDSYVIPALQFAEVSDR	True
peg.1401	TNNFNFADNNGNEIM[15.9949]EVR	True
peg.1401	TNNFNFADNNGNEIM[15.9949]EVR	True
peg.1401	TNNFNFADNNGNEIM[15.9949]EVR	True
peg.1401	TNNFNFADNNGNEIM[15.9949]EVR	True
peg.1401	TNNFNFADNNGNEIM[15.9949]EVR	True

peg.1401	TNNFNFADNNGNEIMEVR	True
peg.1401	TNNFNFADNNGNEIMEVR	True
peg.1401	TNNFNFADNNGNEIMEVR	True
peg.1401	SFLEDTPDQATDGSIK	True
peg.1401	SFLEDTPDQATDGSIK	True
peg.1401	SFLEDTPDQATDGSIK	True
peg.1401	SFLEDTPDQATDGSIK	True
peg.1401	SFLEDTPDQATDGSIK	True
peg.1401	SFLEDTPDQATDGSIK	True
peg.1401	SFLEDTPDQATDGSIK	True
peg.1401	SFLEDTPDQATDGSIK	True
peg.1401	SFLEDTPDQATDGSIK	True
peg.1401	SFLEDTPDQATDGSIK	True
peg.1401	SFLEDTPDQATDGSIK	True
peg.1401	SFLEDTPDQATDGSIK	True
peg.1401	SFLEDTPDQATDGSIK	True
peg.1401	SFLEDTPDQATDGSIK	True
peg.1401	SFLEDTPDQATDGSIK	True
peg.1401	SFLEDTPDQ[0.9840]ATDGSIK	True
peg.1401	SFLEDTPDQ[0.9840]ATDGSIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True

peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	GIIEVSDVFDAYIK	True
peg.1401	IEYDSPTTENIIVGFAPDNTKDFYSK	True
peg.1401	IEYDSPTTENIIVGFAPDNTKDFYSK	True
peg.1401	IEYDSPTTENIIVGFAPDNTKDFYSK	True
peg.1401	IEYDSPTTENIIVGFAPDNTKDFYSK	True
peg.1401	IEYDSPTTENIIVGFAPDNTKDFYSK	True

peg.1401	NM[15.9949]CNLYVFPFAEAWSLMR	True
peg.1401	NM[15.9949]CNLYVFPFAEAWSLMR	True
peg.1401	NM[15.9949]CNLYVFPFAEAWSLMR	True
peg.1401	NMCNLYVFPFAEAWSLM[15.9949]R	True
peg.1401	IEYDSPTTENIIVGFAPDNTK	True
peg.1401	IEYDSPTTENIIVGFAPDNTK	True
peg.1401	IEYDSPTTENIIVGFAPDNTK	True
peg.1401	IEYDSPTTENIIVGFAPDNTK	True
peg.1401	IEYDSPTTENIIVGFAPDNTK	True
peg.1401	IEYDSPTTENIIVGFAPDNTK	True
peg.1401	IEYDSPTTENIIVGFAPDNTK	True
peg.1401	IEYDSPTTENIIVGFAPDNTK	True
peg.1401	IEYDSPTTENIIVGFAPDNTK	True
peg.1401	IEYDSPTTENIIVGFAPDNTK	True
peg.1401	TGTIPPNYLAYDGYYIR	True
peg.1401	TGTIPPNYLAYDGYYIR	True
peg.1401	TGTIPPNYLAYDGYYIR	True
peg.1401	TGTIPPNYLAYDGYYIR	True
peg.1401	TGTIPPNYLAYDGYYIR	True
peg.1401	TGTIPPNYLAYDGYYIR	True
peg.1401	TGTIPPNYLAYDGYYIR	True
peg.1401	TGTIPPNYLAYDGYYIR	True
peg.1401	TGTIPPNYLAYDGYYIR	True
peg.1401	TGTIPPNYLAYDGYYIR	True
peg.1401	TGTIPPNYLAYDGYYIR	True
peg.1401	TGTIPPNYLAYDGYYIR	True
peg.1401	TGTIPPNYLAYDGYYIR	True
peg.1401	TGTIPPNYLAYDGYYIR	True
peg.1401	TGTIPPNYLAYDGYYIR	True
peg.1401	TGTIPPNYLAYDGYYIR	True
peg.1401	TGTIPPNYLAYDGYYIR	True
peg.1401	TGTIPPNYLAYDGYYIR	True
peg.1401	TGTIPPNYLAYDGYYIR	True
peg.1401	TGTIPPN[0.9840]YLAYDGYYIR	True
peg.1401	TQTFYQNPNNEPIAPR	True
peg.1401	TQTFYQNPNNEPIAPR	True
peg.1401	VLSLIFPGSQPATMEK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True

peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	DIINQILTAPAPADLFFK	True
peg.1401	TEVETLIN[0.9840]QK	True
peg.1401	FTQWFQSTLYGWNIK	True
peg.1401	FTQWFQSTLYGWNIK	True
peg.1401	M[15.9949]YTEEFGR	True
peg.1401	MYTEEFGR	True
peg.1401	DGLTFR	True
peg.1401	DGLTFR	True
peg.1401	DGLTFR	True
peg.1401	DGLTFR	True
peg.1401	DGLTFR	True
peg.1401	DGLTFR	True
peg.1401	LPAGIR	True
peg.1401	LPAGIR	True
peg.1401	LPAGIR	True
peg.1401	LPQFEVQTYEGVSIALFTQMCTLHLTLLK	True
peg.1401	LPQFEVQTYEGVSIALFTQMCTLHLTLLK	True
peg.1401	NMCNLYVFPFAEAWSLMR	True
peg.1401	NM[15.9949]CNLYVFPFAEAWSLM[15.9949]R	True

peg.1401	GVSLAYNHDLTTLTYNRIEYDSPTTENIIVGFAPDNTK	True
peg.1401	DGILAGSAWGFTQ[0.9840]ADVDSFIK	True
peg.1401	DGILAGSAWGFTQ[0.9840]ADVDSFIK	True
peg.1401	DIINQILTAPAPADLFFKNADINVK	True
peg.1401	DIINQILTAPAPADLFFKN[0.9840]ADINVK	True
peg.1401	GIIEVSDVFDAYIKQPGFTPATAK	True
peg.1401	GIIEVSDVFDAYIKQPGFTPATAK	True
peg.1401	LPAGIR	True
peg.1401	DGLTFR	True
peg.1401	FTQWFQSTLYGWNIK	True
peg.1401	FTQWFQSTLYGWNIK	True
peg.1401	GYFLNLSGAIIQR	True
peg.1401	VRTEVETLINQK	True
peg.1401	VRVPYRLPAGIR	True
peg.1401	VPYRLPAGIR	True
peg.1401	ESAFTTQINPLLK	True

Table S2.4 Peptides identified to Mpp60Ba3 protein classified as unique (true) or non-unique (false).

CDS id	Peptide sequence	unique
peg.3553	DGQILNVYDNR	True
peg.3553	DGQILNVYDNR	True
peg.3553	DGQILNVYDNR	True
peg.3553	DGQILNVYDNR	True
peg.3553	DGQILNVYDNR	True
peg.3553	TEMISPASR	True
peg.3553	TEMISPASR	True
peg.3553	YTETPLDR	True
peg.3553	YTETPLDR	True
peg.3553	YTETPLDR	True
peg.3553	YTETPLDR	True
peg.3553	GLYAFIR	True
peg.3553	GLYAFIR	True
peg.3553	M[15.9949]EITDIVLK	True
peg.3553	M[15.9949]EITDIVLK	True
peg.3553	M[15.9949]EITDIVLK	True
peg.3553	M[15.9949]EITDIVLK	True
peg.3553	M[15.9949]EITDIVLK	True
peg.3553	M[15.9949]EITDIVLK	True
peg.3553	SSQTQLHTIK	True
peg.3553	SSQTQLHTIK	True
peg.3553	SSQTQLHTIK	True
peg.3553	GIVKPVR	True
peg.3553	GIVKPVR	True
peg.3553	YYDDYANMFFSYIFQSK	True
peg.3553	YYDDYANMFFSYIFQSK	True
peg.3553	YYDDYANMFFSYIFQSK	True
peg.3553	AIYEYELN[0.9840]DTVTIPETK	True
peg.3553	AIYEYELN[0.9840]DTVTIPETK	True
peg.3553	AIYEYELNDTVTIPETK	True
peg.3553	AIYEYELNDTVTIPETK	True
peg.3553	AIYEYELNDTVTIPETK	True
peg.3553	SGGSNGSLNLSGFGYSDLYK	True
peg.3553	VFQTTPIPIASALTITENR	True
peg.3553	VFQTTPIPIASALTITENR	True
peg.3553	MESVTNTTVHGFK	True
peg.3553	MESVTNTTVHGFK	True
peg.3553	KMESVTNTTVHGFK	True
peg.3553	KMESVTNTTVHGFK	True
peg.3553	VPM[15.9949]ILNSNLIGK	True
peg.3553	VPM[15.9949]ILNSNLIGK	True

peg.3553	VPM[15.9949]ILNSNLIGK	True
peg.3553	VPMILNSNLIGK	True
peg.3553	VPMILNSNLIGK	True
peg.3553	VPMILNSNLIGK	True
peg.3553	TWDSNLIHLR	True
peg.3553	TWDSNLIHLR	True
peg.3553	LANQSWPGKPIVFK	True
peg.3553	MEITDIVLK	True
peg.3553	MEITDIVLK	True
peg.3553	MEITDIVLK	True
peg.3553	MEITDIVLK	True
peg.3553	MEITDIVLK	True
peg.3553	MEITDIVLK	True
peg.3553	MEITDIVLK	True
peg.3553	MEITDIVLK	True
peg.3553	MEITDIVLK	True
peg.3553	TEM[15.9949]ISPASR	True
peg.3553	TEM[15.9949]ISPASR	True
peg.3553	TEM[15.9949]ISPASR	True
peg.3553	RYYDDYANMFFSYIFQSK	True
peg.3553	KM[15.9949]ESVTNTTVHGFK	True
peg.3553	KM[15.9949]ESVTNTTVHGFK	True
peg.3553	KM[15.9949]ESVTNTTVHGFK	True
peg.3553	YSSPGKTWDSNLIHLRDGQILNVYDNR	True
peg.3553	YSSPGKTWDSNLIHLRDGQILNVYDNR	True
peg.3553	M[15.9949]ESVTNTTVHGFK	True
peg.3553	TWEITENVSVASHTSLTSQLIIMQADIR	True
peg.3553	YYDDYANMFFSYIFQSKTSGRTEMISPASR	True
peg.3553	IYDFIEWDYVTNQDGIPYTLFDK	True
peg.3553	IYDFIEWDYVTNQDGIPYTLFDK	True
peg.3553	LANQSWPGKPIVFK	True
peg.3553	AIYEYELNDTVTIPETK	True
peg.3553	AIYEYELN[0.9840]DTVTIPETK	True
peg.3553	AIYEYELN[0.9840]DTVTIPETK	True
peg.3553	TSGRTEM[15.9949]ISPASR	True
peg.3553	SGGSNGSLNLSGFGYSDLYKGLYAFIRYTETPLDR	True
peg.3553	SSQTQLHTIK	True
peg.3553	SSQTQLHTIK	True
peg.3553	TWDSNLIHLR	True

Table S2.5 Peptides identified to Cry10Aa4 protein classified as unique (true) or non-unique (false).

CDS id	Peptide sequence	unique
peg.6260	RTDSYM[15.9949]IPK	True
peg.6260	VFPFYR	True
peg.6260	GPGHTGGDLVILK	True
peg.6260	NKEYGHTLSYIK	True
peg.6260	TTVFLTGIDTISVELPSTTSR	True
peg.6260	YSSNSTIENNYKR	True
peg.6260	NKPIDKVEIVR	True
peg.6260	EYGHTLSYIK	True
peg.6260	NKPIDKVEIVR	True
peg.6260	NQLDKYQEFFDK	True
peg.6260	NKPIDKVEIVR	True
peg.6260	NEYEIFNAPSN[0.9840]GFSK	False
peg.6260	YPLANKPNQPLK	False
peg.6260	YPLANKPNQPLK	False
peg.6260	YPLANKPNQPLK	False
peg.6260	NQLDKYQEFFDKWEPAR	True
peg.6260	NQLDKYQEFFDKWEPAR	True
peg.6260	NQLDKYQEFFDKWEPAR	True
peg.6260	YQEFFDKWEPAR	True
peg.6260	YQEFFDKWEPAR	True
peg.6260	YQEFFDKWEPAR	True
peg.6260	NIIWGAVHGNIISQDTSK	True
peg.6260	NIIWGAVHGNIISQDTSK	True
peg.6260	AVHDLFTTLEPIIDK	True
peg.6260	AVHDLFTTLEPIIDK	True
peg.6260	TITELENGLTR	True
peg.6260	TITELENGLTR	True
peg.6260	TITELENGLTR	True
peg.6260	TDNYIFSVVR	True
peg.6260	TDNYIFSVVR	True
peg.6260	QNPNATDLTYADFGYVTFPR	True
peg.6260	QNPNATDLTYADFGYVTFPR	True
peg.6260	QNPNATDLTYADFGYVTFPR	True
peg.6260	QNPNATDLTYADFGYVTFPR	True
peg.6260	QNPNATDLTYADFGYVTFPR	True
peg.6260	AVHDLFTTLEPIIDKDLDMLK	True
peg.6260	AVHDLFTTLEPIIDKDLDMLK	True
peg.6260	AVHDLFTTLEPIIDKDLDMLK	True
peg.6260	AVHDLFTTLEPIIDKDLDMLK	True
peg.6260	AVHDLFTTLEPIIDKDLDMLK	True
peg.6260	AVHDLFTTLEPIIDKDLDMLK	True

peg.6260	AVHDLFTTLEPIIDKDLDMLK	True
peg.6260	AVHDLFTTLEPIIDKDLDMLK	True
peg.6260	AVHDLFTTLEPIIDKDLDMLK	True
peg.6260	AVHDLFTTLEPIIDKDLDMLK	True
peg.6260	NIINVLTSIVTPIK	True
peg.6260	TNTNATWNMYNTYR	True
peg.6260	NPTLFTWINQGR	True
peg.6260	YSSNSTIENNYK	True
peg.6260	EVYTNVNSDTFR	True
peg.6260	EVYTNVNSDTFR	True
peg.6260	NKEYGHTLSYIK	True
peg.6260	TTVFLTGIDTISVELPSTTSR	True
peg.6260	GPGHTGGDLVILK	True
peg.6260	TITELEN[0.9840]GLTR	True
peg.6260	TDSYMIPK	True
peg.6260	TDSYM[15.9949]IPK	True
peg.6260	EYGHTLSYIK	True
peg.6260	DILDPYDIFSFTGNQMAFTHTNDDR	True
peg.6260	DILDPYDIFSFTGNQM[15.9949]AFTHTNDDR	True
peg.6260	AVHDLFTTLEPIIDKDLDM[15.9949]LK	True
peg.6260	AVHDLFTTLEPIIDKDLDM[15.9949]LK	True
peg.6260	AVHDLFTTLEPIIDKDLDM[15.9949]LK	True
peg.6260	AVHDLFTTLEPIIDKDLDM[15.9949]LK	True
peg.6260	AVHDLFTTLEPIIDKDLDM[15.9949]LK	True
peg.6260	AVHDLFTTLEPIIDKDLDM[15.9949]LK	True
peg.6260	AVHDLFTTLEPIIDKDLDM[15.9949]LK	True
peg.6260	IPTLPAYAQIATWHLNLLK	True

Table S2.6 Peptides identified to Cyt1Aa5 protein classified as unique (true) or non-unique (false).

CDS id	Peptide sequence	unique
peg.5774	TPQSTAR	True
peg.5774	M[15.9949]ENLNHCPLEDIKVNPWK	True
peg.5774	M[15.9949]ENLNHCPLEDIKVNPWK	True
peg.5774	M[15.9949]ENLNHCPLEDIKVNPWK	True
peg.5774	M[15.9949]ENLNHCPLEDIKVNPWK	True
peg.5774	TVLGVALSGSVIDQLTAAVTNTFTNLNTQK	True
peg.5774	TVLGVALSGSVIDQLTAAVTNTFTNLNTQK	True
peg.5774	M[15.9949]ENLNHCPLEDIK	True
peg.5774	MENLNHCPLEDIK	True
peg.5774	M[15.9949]ENLNHCPLEDIK	True
peg.5774	MENLNHCPLEDIK	True
peg.5774	MENLNHCPLEDIK	True
peg.5774	M[15.9949]ENLNHCPLEDIK	True
peg.5774	M[15.9949]ENLNHCPLEDIK	True
peg.5774	MENLNHCPLEDIK	True
peg.5774	MENLNHCPLEDIK	True
peg.5774	M[15.9949]ENLNHCPLEDIK	True
peg.5774	M[15.9949]ENLNHCPLEDIK	True
peg.5774	GLEIANTITPMGAVVSYVDQNVTQTNNQVSVMINK	True
peg.5774	GLEIANTITPMGAVVSYVDQNVTQTNNQVSVMINK	True
peg.5774	GLEIANTITPMGAVVSYVDQNVTQTNNQVSVMINK	True
peg.5774	ETANQTNYTYNVLFAIQNAQTGGVMYCVPVGFEIK	True
peg.5774	ETANQTNYTYNVLFAIQNAQTGGVMYCVPVGFEIK	True
peg.5774	FAQPLVSSSQYPIADLTSAINGTL	True
peg.5774	FAQPLVSSSQYPIADLTSAINGTL	True
peg.5774	FAQPLVSSSQYPIADLTSAINGTL	True
peg.5774	FAQPLVSSSQYPIADLTSAINGTL	True
peg.5774	FAQPLVSSSQYPIADLTSAINGTL	True
peg.5774	FAQPLVSSSQYPIADLTSAINGTL	True
peg.5774	FAQPLVSSSQYPIADLTSAINGTL	True
peg.5774	FAQPLVSSSQYPIADLTSAINGTL	True
peg.5774	FAQPLVSSSQYPIADLTSAINGTL	True
peg.5774	FAQPLVSSSQYPIADLTSAINGTL	True
peg.5774	FAQPLVSSSQYPIADLTSAINGTL	True
peg.5774	FAQPLVSSSQYPIADLTSAINGTL	True
peg.5774	FAQPLVSSSQYPIADLTSAINGTL	True
peg.5774	FAQPLVSSSQYPIADLTSAINGTL	True
peg.5774	FAQPLVSSSQYPIADLTSAINGTL	True
peg.5774	FAQPLVSSSQYPIADLTSAINGTL	True
peg.5774	FAQPLVSSSQYPIADLTSAINGTL	True
peg.5774	FAQPLVSSSQYPIADLTSAINGTL	True

peg.5774	FAQPLVSSSQYPIADLTSAINGTL	True
peg.5774	FAQPLVSSSQYPIADLTSAINGTL	True
peg.5774	FAQPLVSSSQYPIADLTSAINGTL	True
peg.5774	FAQPLVSSSQYPIADLTSAINGTL	True
peg.5774	FAQPLVSSSQYPIADLTSAINGTL	True
peg.5774	FAQPLVSSSQYPIADLTSAINGTL	True
peg.5774	FAQPLVSSSQYPIADLTSAINGTL	True
peg.5774	FAQPLVSSSQYPIADLTSAIN[0.9840]GTL	True
peg.5774	FAQPLVSSSQYPIADLTSAIN[0.9840]GTL	True
peg.5774	FAQPLVSSSQYPIADLTSAIN[0.9840]GTL	True
peg.5774	FAQPLVSSSQYPIADLTSAIN[0.9840]GTL	True
peg.5774	FAQPLVSSSQYPIADLTSAIN[0.9840]GTL	True
peg.5774	FAQPLVSSSQ[0.9840]YPIADLTSAIN[0.9840]GTL	True
peg.5774	N[0.9840]EAWIFWGK	True
peg.5774	N[0.9840]EAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True

peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	NEAWIFWGK	True
peg.5774	MENLNHCPLEDIKVNPWK	True
peg.5774	MENLNHCPLEDIKVNPWK	True
peg.5774	MENLNHCPLEDIKVNPWK	True

peg.5774	MENLNHCPLEDIKVNPWK	True
peg.5774	MENLNHCPLEDIKVNPWK	True
peg.5774	MENLNHCPLEDIKVNPWK	True
peg.5774	MENLNHCPLEDIK	True
peg.5774	MENLNHCPLEDIK	True
peg.5774	MENLNHCPLEDIK	True
peg.5774	M[15.9949]ENLNHCPLEDIK	True
peg.5774	M[15.9949]ENLNHCPLEDIK	True

Table S2.7 Peptides identified to Cyt2Ba13 protein classified as unique (true) or non-unique (false).

CDS id	Peptide sequence	unique
peg.3388	QTSYFYK	True
peg.3388	QTSYFYK	True
peg.3388	QTSYFYK	True
peg.3388	QQLLFITIK	True
peg.3388	QQLLFITIK	True
peg.3388	QQLLFITIK	True
peg.3388	ALTVVQALDSYNAPIIDVFNVR	True
peg.3388	ALTVVQALDSYNAPIIDVFNVR	True
peg.3388	FMAILPIAFEITVDVQK	True
peg.3388	FMAILPIAFEITVDVQK	True
peg.3388	ILFSIQNEDTGR	True
peg.3388	FM[15.9949]AILPIAFEITVDVQK	True
peg.3388	NYSLHRPNHNILQNLNVNPIK	True
peg.3388	NYSLHRPNHNILQNLNVNPIK	True

Table S2.8 Peptides identified to Cyt1Ca1 protein classified as unique (true) or non-unique (false)

CDS id	Peptide sequence	unique
peg.7343	KYEYLQIYSSGK	True
peg.7343	MLQHLFLDLR	True
peg.7343	M[15.9949]LQHLFLDLR	True
peg.7343	YEYLQIYSSGK	True
peg.7343	MLQHLFLDLR	True
peg.7343	M[15.9949]LQHLFLDLR	True
peg.7343	KYEYLQIYSSGK	True
peg.7343	LEGQGMYVGLVEDNNQTFLCWR	True

5.1 Appendice B- Supplemental material of Chapter 3

Table S1 Toxicity (larval mortality) of *Bt* strains (UFT038 and HD-1) against six lepidopteran species and two Cry1-*Bt*-resistant populations of *Spodoptera frugiperda*.

Bt strain	Insect	LC_{50} (95% CL)[1] (10^6 spores + crystals/cm^2)	ResR[2] (95% CL) [Among insect species within *Bt* strain]	ToxR[3] (95% CL) [Between *Bt* strains within insect species or population]	CRR[4] (95% CL) [Between a resistant and susceptible FAW population]	Slope ± SE	χ^{2} [5]	*P*	*N*
UFT038	*S. cosmioides*	6.8 (2.8 - 12.8)	1	2.6 (0.6 - 11.2)^{ns}	-	0.62 ± 0.08	1.4	0.93	504
	S. frugiperda S_Bt	18.3 (0.01 - 96.3)	2.7 (0.4 - 16.9)^{ns}	1	1	0.73 ± 0.15	9.0	0.11	638
	H. zea	25.2 (9.9 - 51.3)	3.7 (1.5 - 9.1)*	1	-	1.25 ± 0.13	5.2	0.26	448
	H. armigera	36.5 (19.7 - 60.4)	5.4 (2.3 - 12.5)*	1	-	1.17 ± 0.12	4.0	0.40	448
	C. includens	50.3 (27.6 - 98.8)	8.0 (3.4 - 18.9)*	1	-	1.53 ± 0.15	6.8	0.15	448
	S. frugiperda R_Cry1+2Ab	208.9 (93.1 - 373.3)	31.3 (13.0 - 71.4)*	1	11.5 (2.0 - 62.5)*	0.99 ± 0.09	11.2	0.26	689
	S. eridania	218.1 (102.5 - 306.1)	32.3 (14.7 - 71.4)*	1	-	2.75 ± 0.47	9.9	0.08	512
	S. frugiperda R_Cry1F	221.2 (86.4 - 468.1)	32.3 (13.5 - 76.9)*	1	12.0 (2.1 - 66.7)*	1.00 ± 0.11	9.6	0.09	697
HD-1	*C. includens*	2.3 (1.8 - 3.0)	1	22.2 (14.5 - 33.7)*	-	1.54 ± 0.14	1.4	0.93	512
	H. armigera	14.6 (2.9 - 40.0)	6.2 (3.1 - 12.2)*	2.5 (1.2 - 5.0)*	-	0.92 ± 0.14	10.0	0.08	512
	H. zea	16.2 (7.9 - 27.8)	6.8 (3.4 - 13.7)*	1.5 (0.7 - 3.3)^{ns}	-	0.93 ± 0.15	3.6	0.61	512
	S. cosmioides	17.5 (5.1 - 80.1)	7.4 (2.0 - 27.8)*	1	-	0.36 ± 0.08	2.8	0.74	511
	S. eridania	57.3 (37.9 - 84.5)	24.4 (16.1 - 35.7)*	3.8 (2.8 - 5.1)*	-	1.87 ± 0.16	9.9	0.08	510
	S. frugiperda S_*Bt*	56.7 (19.2 - 96.4)	23.8 (14.3 - 40.0)*	3.1 (0.6 - 17.3)^{ns}	1	1.73 ± 0.22	7.8	0.17	574
	S. frugiperda R_Cry1F	43.7 (22.1 - 80.6)	18.5 (10.8 - 32.3)*	5.1 (2.7 - 9.5)*	0.8 (0.4 - 1.4)^{ns}	0.83 ± 0.07	8.2	0.15	764
	S. frugiperda R_Cry1+2Ab	130.4 (73.7 - 212.7)	55.6 (33.3 - 90.9)*	1.6 (0.9 - 2.8)^{ns}	2.3 (1.4 - 3.9)*	1.05 ± 0.13	9.6	0.21	681

* Significant ($P > 0.05$), since the confidence interval did not include the value 1 calculated according to Robertson et al. (2017)

^{ns} Not significant ($P > 0.05$), given that the confidence interval include the value 1 (Robertson et al. 2017).

[1] The toxin concentration (10^6 spores+crystals/cm^2) causing 50% (CL_{50}) ou 90% (CL_{90}) larval mortality and in parentheses, their respective confidence limits with a 95% probability (95% CI).

[2] Resistance ratio: ratio of the LC_{50} estimate of species most tolerant to HD-1 or UFT038.

[3] Toxicity ratio: ratio of the LC_{50} estimate between HD-1 and UFT38.

[4] Cross-resistance ratio: ratio of the LC_{50} estimate of species resistance to HD-1 or UFT038.

[5] Chi-Square ($P > 0.05$).

Table S2 Toxicity (growth inhibition) of *Bt* strains against six lepidopteran species and two Cry1-Bt-resistant populations of *Spodoptera frugiperda*.

Bt strain	Insect species or population	EC_{50} (95% CL)[1] (10^6 spores + crystals/cm^2)	ResR[2] (95% CL) [Among insect species within *Bt* strain]	ToxR[3] (95% CL) [Between *Bt* strains within insect species or population]	CRR[4] (95% CL) [Between a resistant and susceptible FAW population]	Slope ± SE	χ^2 [5]	*P*	*N*
UFT038	*S. cosmioides*	0.2 (0.001 - 1.1)	1	27.9 (3.4 - 228.3)*	-	0.53 ± 0.07	5.1	0.4	504
	H. zea	1.6 (0.3 - 3.4)	9.3 (1.1 - 83.3)	1	-	0.94 ± 0.14	4.4	0.35	448
	H. armigera	1.9 (0.8 - 3.0)	10.6 (1.3 - 90.9)	1	-	1.18 ± 0.18	2.2	0.69	448
	S. frugiperda S_Bt	3.9 (1.7 - 7.2)	22.7 (2.8 - 200)	2.7 (1.8 - 5.4)*	1	0.93 ± 0.09	9.7	0.2	703
	S. frugiperda R_Cry1F	23.7 (14.2 - 37.1)	136 (17.4 - 1064)	1	6.0 (3.3 - 11.0)*	0.91 ± 0.07	11.4	0.18	713
	S. eridania	25.6 (15.4 - 38.0)	145 (18.4 - 1173)	1	-	1.40 ± 0.16	4.1	0.54	512
	C. includens	38.6 (30.2 - 50.5)	222 (28.9 - 1697)	1	-	2.00 ± 0.13	5.4	0.25	448
	S. frugiperda R_Cry1+2Ab	47.8 (23.3 - 83.0)	274 (34.8 - 2164.1)	1	12.0 (6.4 - 22.7)*	1.01 ± 0.08	12.6	0.08	689
HD-1	*C. includens*	0.6 (0.4 - 0.8)	1	69.6 (51.5 - 94)*	-	1.54 ± 0.14	6.5	0.26	512
	H. armigera	0.8 (0.5 - 1.2)	1.5 (1.0 - 2.2)	2.3 (1.2 - 4.6)*	-	1.21 ± 0.12	6.7	0.24	512
	H. zea	0.9 (0.6 - 1.2)	1.6 (1.0 - 2.5)	1.8 (0.8 - 4.2)ns	-	0.98 ± 0.09	3.6	0.61	512
	S. cosmioides	4.9 (1.8 - 12.7)	8.8 (4.8 - 16.1)	1	-	0.53 ± 0.07	7.6	0.18	479
	S. frugiperda R_Cry1F	8.1 (5.6 - 11.5)	14.7 (9.4 - 22.7)	2.9 (1.9 - 4.8)*	0.6 (0.4 - 1.0)[ns]	0.75 ± 0.08	5.3	0.73	763
	S. frugiperda S_Bt	12.5 (9.6 - 17.2)	22.7 (16.1 - 31.3)	1	1	2.97 ± 0.45	0.3	0.99	574
	S. eridania	13.7 (7.9 - 20.2)	25.0 (17.5 - 34.5)	1.9 (1.1 - 3.1)*	-	1.97 ± 0.19	11.1	0.05	510
	S. frugiperda R_Cry1+2Ab	23.9 (14.1 - 37.1)	43.5 (29.4 - 62.5)	2.0 (1.2 - 3.2)*	1.9 (1.3 - 2.8)*	1.11 ± 0.09	10.1	0.12	630

* Significant ($P > 0.05$), since the confidence interval did not include the value 1 calculated according to Robertson et al. (2017)

[ns] Not significant ($P > 0.05$), given that the confidence interval include the value 1 (Robertson et al. 2017).

[1] The toxin concentration (10^6 spores+crystals/cm^2) causing 50% (CL_{50}) ou 90% (CL_{90}) larval mortality and in parentheses, their respective confidence limits with a 95% probability (95% CI).

[2] Resistance ratio: ratio of the EC_{50} estimate of species most tolerant to HD-1 or UFT038.

[3] Toxicity ratio: ratio of the EC_{50} estimate between HD-1 and UFT38.

[4] Cross-resistance ratio: ratio of the LC_{50} estimate of species resistance to HD-1 or UFT038.

[5] Chi-Square ($P > 0.05$).

6 ANNEXES

6.1 Annexe A: Chapter I published in Scientific Reports.

www.nature.com/scientificreports

SCIENTIFIC REPORTS
nature research

OPEN

Comparative genomic analysis and mosquito larvicidal activity of four *Bacillus thuringiensis* serovar *israelensis* strains

Giselly B. Alves[1], Fernando L. Melo[2], Eugenio E. Oliveira[3], Khalid Haddi[4], Lara T. M. Costa[1], Marcelo L. Dias[1], Fabrício S. Campos[1], Eliseu J. G. Pereira[3], Roberto F. T. Corrêa[1], Sergio D. Ascêncio[5], Gil R. Santos[1], Guy Smagghe[6], Bergmann M. Ribeiro[2] & Raimundo W. S. Aguiar[1*]

Bacillus thuringiensis **serovar** ***israelensis*** **(*Bti*) is used to control insect vectors of human and animal diseases. In the present study, the toxicity of four strains of *Bti*, named T0124, T0131, T0137, and T0139, toward *Aedes aegypti* and *Culex quinquefasciatus* larvae was analyzed. The T0131 strain showed the highest larvicidal activity against *A. aegypti* (LC_{50} = 0.015 µg/ml) and *C. quinquefasciatus* larvae (LC_{50} = 0.035 µg/ml) when compared to the other strains. Furthermore, the genomic sequences of the four strains were obtained and compared. These *Bti* strains had chromosomes sizes of approximately 5.4 Mb with GC contents of ~35% and 5472–5477 putative coding regions. Three small plasmids (5.4, 6.8, and 7.6 kb) and three large plasmids (127, 235, and 359 kb) were found in the extrachromosomal content of all four strains. The SNP-based phylogeny revealed close relationship among isolates from this study and other *Bti* isolates, and SNPs analysis of the plasmids 127 kb did not reveal any mutations in δ-endotoxins genes. This newly acquired sequence data for these *Bti* strains may be useful in the search for novel insecticidal toxins to improve existing ones or develop new strategies for the biological control of important insect vectors of human and animal diseases.**

During sporulation, the gram-positive bacterium *Bacillus thuringiensis* (*Bt*) produces crystalline inclusions consisting of δ-endotoxins (Cry or Cyt proteins) with insecticidal activity[1]. Genomic analysis has contributed to the identification of new genes coding for toxins that are active against different insect species including orders such as Lepidoptera, Diptera[2–7], and Coleoptera[8]. Proteins with nematicidal[9–11] and molluscicidal[12] activities have also been described. In addition, genome sequencing of *Bt* strains with diverse ecological functions has been conducted, including a endophytic strain with potential utility in the biocontrol of phytopathogens[13].

Sequencing of complete *Bt* genomes has allowed structural and functional analysis of new plasmids that enhance our knowledge of the pathogenic properties of *Bt* in targeting organisms[14–17]. One study reported the plasmid sequence of a *Bacillus thuringiensis* serovar *israelensis* (*Bti*) strain, and revealed that it may produce up to seven crystal-forming toxins, named Cry4A, Cry4B, Cry10A, Cry11A, Cyt1A, Cyt2Ba, and Cyt1Ca, which are all encoded by genes found in a single 127923 bp plasmid called pBtoxis[18]. The average size of the complete genome sequences of *Bti* is 6.1 Mb, with ~35% GC content of the chromosomal DNA and an average of 6132 coding sequences[19,20]. Genome sequences of seven *Bti* isolates have been reported so far[18,21–24].

In this study, we sequenced the genomes of four *Bti* strains, specifically T0124, T0131, T0137, and T0139 that were collected from the soil of the Tocantins state in Brazil and determined their larvicidal activity against larvae of two important mosquito species of *A. aegypti* and *C. quinquefasciatus*. Then, to better characterize these strains, we performed comparative and phylogenetic analyses among their different genomes and compared the

[1]Departamento de Biotecnologia, Universidade Federal de Tocantins, Gurupi, TO, 77413-070, Brazil. [2]Departamento de Biologia Celular, Universidade de Brasília, Brasília, DF, 70910-900, Brazil. [3]Departamento de Entomologia, Universidade Federal de Viçosa, Viçosa, MG, 36570-900, Brazil. [4]Departamento de Entomologia, Universidade Federal de Lavras, Lavras, MG, 37200-900, Brazil. [5]Rede de Biodiversidade e Biotecnologia da Amazônia Legal (Rede Bionorte), Universidade Federal do Tocantins, Palmas, TO, 77413-070, Brazil. [6]Department of Plants and Crops, Ghent University, 9000, Ghent, Belgium. *email: rwsa@uft.edu.br

6.2 Annexe B: Chapter II submitted in Processes.

[Processes] Manuscript ID: processes-2273325 - Submission Received

2 mensagens

Editorial Office <processes@mdpi.com> 24 de fevereiro de 2023 às 05:43

Dear Dr. Aguiar,

Thank you very much for uploading the following manuscript to the MDPI submission system. One of our editors will be in touch with you soon.

Journal name: Processes
Manuscript ID: processes-2273325
Type of manuscript: Article
Title: Isolation, genomic and proteomic characterization of a novel Neotropical strain of *Bacillus thuringiensis* with mosquitocidal activities
Authors: Giselly Batista Alves, Marcelo Leite Dias, Eugenio Eduardo De Oliveira, Bergmann Morais Ribeiro, Raimundo Wagner de Souza Souza Aguiar *
Received: 24 February 2023
E-mails: gybiotec@gmail.com, marcelo_70dias@hotmail.com, eugenio@ufv.br, bergmann.ribeiro@gmail.com, rwsa@uft.edu.br

You can follow progress of your manuscript at the following link (login required):
https://susy.mdpi.com/user/manuscripts/review_info/94e103a0ebebe979aec3b27f73b5ac7f

The following points were confirmed during submission:

1. Processes is an open access journal with publishing fees of 2000 CHF for an accepted paper (see https://www.mdpi.com/about/apc/ for details). This manuscript, if accepted, will be published under an open access Creative Commons CC BY license (https://creativecommons.org/licenses/by/4.0/), and I agree to pay the Article Processing Charges as described on the journal webpage (https://www.mdpi.com/journal/processes/apc). See https://www.mdpi.com/about/openaccess for more information about open access publishing.

Please note that you may be entitled to a discount if you have previously received a discount code or if your institute is participating in the MDPI Institutional Open Access Program (IOAP), for more information see https://www.mdpi.com/about/ioap. If you have been granted any other special discounts for your submission, please contact the Processes editorial office.

2. I understand that:

a. If previously published material is reproduced in my manuscript, I will provide proof that I have obtained the necessary copyright permission. (Please refer to the Rights & Permissions website: https://www.mdpi.com/authors/rights).

b. My manuscript is submitted on the understanding that it has not been published in or submitted to another peer-reviewed journal. Exceptions to this rule are papers containing material disclosed at conferences. I confirm that I will inform the journal editorial office if this is the case for my manuscript. I confirm that all authors are familiar with and agree with submission of the contents of the manuscript. The journal editorial office reserves the right to contact all authors to confirm this in case of doubt. I will provide email addresses for all authors and an institutional e-mail address for at least one of the co-authors, and specify the name, address and e-mail for invoicing purposes.

If you have any questions, please do not hesitate to contact the Processes editorial office at processes@mdpi.com

Kind regards,
Processes Editorial Office
St. Alban-Anlage 66, 4052 Basel, Switzerland
E-Mail: processes@mdpi.com

6.3 Annexe C: Chapter III submitted in Journal of Asia-Pacific Entomology

Confirm co-authorship of submission to Journal of Asia-Pacific Entomology

1 mensagem

Journal of Asia-Pacific Entomology <em@editorialmanager.com> 24 de fevereiro de 2023 às 20:59

This is an automated message.

Journal: Journal of Asia-Pacific Entomology

Title: Neotropical *Bacillus thuringiensis* strain (UFT038): genomic analysis and its potential against multiple lepidopteran pests of soybean

Corresponding Author: doctor Raimundo Wagner Aguiar

Co-Authors: Giselly B. Alves; Timóteo S. O. Lemes; Eliseu J. G. Pereira; Juan L. Jurat-Fuentes; Guy Smagghe; Gil R. Santos; Khalid Haddi; Roberto F. T. Correia; Fernando L. Melo; Luis O. Viteri Jumbo; Eugenio E. Oliveira; Bergmann M. Ribeiro

Manuscript Number:

Dear Alves,

doctor Raimundo Wagner Aguiar submitted this manuscript via Elsevier's online submission system, Editorial Manager, and you have been listed as a Co-Author of this submission.

Elsevier asks Co-Authors to confirm their consent to be listed as Co-Author and track the papers status. In order to confirm your connection to this submission, please click here to confirm your co-authorship: https://www.editorialmanager.com/jape/l.asp?i=114691&l=IXUCS1AB

If you have not yet registered for the journal on Editorial Manager, you will need to create an account to complete this confirmation. Once your account is set up and you have confirmed your status as Co-Author of the submission, you will be able to view and track the status of the submission as it goes through the editorial process by logging in at https://www.editorialmanager.com/jape/

If you did not co-author this submission, please contact the Corresponding Author directly at rwsa@mail.edu.br

Thank you,

Journal of Asia-Pacific Entomology

6.4 Annexe D: published in IV accepted in Archives of Microbiology

Archives of Microbiology (2023) 205:143
https://doi.org/10.1007/s00203-023-03479-y

ORIGINAL PAPER

Genomic–proteomic analysis of a novel *Bacillus thuringiensis* strain: toxicity against two lepidopteran pests, abundance of Cry1Ac5 toxin, and presence of InhA1 virulence factor

Giselly Batista Alves[1] · Eugênio Eduardo de Oliveira[2] · Luis Oswaldo Viteri Jumbo[1,3] · Gil Rodrigues dos Santos[4] · Manoel Mota dos Santos[5] · Marcio Akio Ootani[1] · Bergmann Morais Ribeiro[6] · Raimundo Wagner de Souza Aguiar[1]

Received: 7 January 2023 / Revised: 9 March 2023 / Accepted: 9 March 2023

Abstract
Bacillus thuringiensis (*Bt*) is a biological alternative to the indiscriminate use of chemical insecticides in agriculture. Due to resistance development on insect pests to *Bt* crops, isolating novel *Bt* strains is a strategy for screening new pesticidal proteins or strains containing toxin profile variety that can delay resistance. Besides, the combined genomic and proteomic approaches allow identifying pesticidal proteins and virulence factors accurately. Here, the genome of a novel *Bt* strain (*Bt* TOL651) was sequenced, and the proteins from the spore–crystal mixture were identified by proteomic analysis. Toxicity bioassays with the spore–crystal mixture against larvae of *Diatraea saccharalis* and *Anticarsia gemmatalis*, key pests of sugarcane and soybean, respectively, were performed. The toxicity of *Bt* TOL651 varies with the insect; *A. gemmatalis* (LC_{50} = 1.45 ng cm^{-2}) is more susceptible than *D. saccharalis* (LC_{50} = 73.77 ng cm^{-2}). Phylogenetic analysis of the *gyrB* gene indicates that TOL651 is related to *Bt kenyae* strains. The genomic analysis revealed the presence of *cry1Aa18*, *cry1Ac5*, *cry1Ia44*, and *cry2Aa9* pesticidal genes. Virulence factor genes such as phospholipases (*plcA*, *piplc*), metalloproteases (*inhA*), hemolysins (*cytK*, *hlyIII*, *hblA*, *hblC*, *hblD*), and enterotoxins (*nheA*, *nheB*, *nheC*) were also identified. The combined use of the genomic and proteomic data indicated the expression of Cry1Aa18, Cry1Ac5, and Cry2Aa9 proteins, with Cry1Ac5 being the most abundant. InhA1 also was expressed and may contribute to *Bt* TOL651 pathogenicity. These results provide *Bt* TOL651 as a new tool for the biocontrol of lepidopteran pests.

Keywords Biopesticide · *Bacillus thuringiensis* · Cry proteins · Genome sequence · Proteomic · Toxicity · Lepidopteran pest

Communicated by Yusuf Akhter.

✉ Raimundo Wagner de Souza Aguiar
rwsa@mail.uft.edu.br

[1] Department of Biotechnology, Universidade Federal de Tocantins, Gurupi, TO 77413-070, Brazil

[2] Department of Entomology, Universidade Federal de Viçosa, Viçosa, MG 36570-900, Brazil

[3] Department of Agronomy, Universidad Nacional de Loja (UNL), 110103 Loja, Loja, Ecuador

[4] Laboratory of Phytopathology, Universidade Federal do Tocantins, Gurupi, TO 77410-530, Brazil

[5] Department of Agricultural Science and Technology, Universidade Federal do Tocantins, Gurupi, TO 77413-070, Brazil

[6] Department of Cell Biology, Universidade de Brasília, Brasília, DF 70910-900, Brazil

Introduction

Sugarcane (*Saccharum officinarum*), corn (*Zea mays*), and soybean (*Glycine max* L) are high-value crops and applied to different purposes such as food and biofuel production (de Matos et al. 2020; Heinrichs et al. 2017). The sugarcane borer, *Diatraea saccharalis* (Fabricius, 1794) (Lepidoptera: Crambidae), is a key pest of sugarcane and corn. The velvetbean caterpillar, *Anticarsia gemmatalis* (Hübner 1818) (Lepidoptera: Noctuidae), is a key pest of soybean (Horikoshi et al. 2022; Mendonça et al. 2020). These insects are among Brazil's most relevant Lepidoptera pests causing damage to the raw material used in food and biofuel processing (Silva 1995; Praça et al. 2004; Moscardi et al. 2012; Dinardo-Miranda et al. 2013).

Biocontrol is a safe alternative to reduce the use of chemical insecticides in crop pest management. Entomopathogenic

Published online: 26 March 2023