Research Article

Journal of Molecular Microbiology and Biotechnology

J Mol Microbiol Biotechnol 2017;27:43–54
DOI: 10.1159/000451056

Received: June 13, 2016
Accepted: September 23, 2016
Published online: January 17, 2017

# Complete Sequence and Organization of pFR260, the *Bacillus thuringiensis* INTA Fr7-4 Plasmid Harboring Insecticidal Genes

Laura E. Navas[a, c] Ariel F. Amadio[b, c] Elio M. Ortiz[a, c] Diego H. Sauka[a, c]
Graciela B. Benintende[a] Marcelo F. Berretta[a, c] Rubén O. Zandomeni[a, c]

[a]Instituto de Microbiología y Zoología Agrícola, Instituto Nacional de Tecnología Agropecuaria (INTA), Castelar,
[b]Estación Experimental Agropecuaria Rafaela, Instituto Nacional de Tecnología Agropecuaria (INTA), Rafaela, and
[c]Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, Argentina

**Keywords**
*Bacillus thuringiensis* · Plasmid · Insecticidal genes

**Abstract**
We report the complete sequence and analysis of pFR260, a novel megaplasmid of 260,595 bp from the *Bacillus thuringiensis* strain INTA Fr7-4 isolated in Argentina. It carries 7 insecticidal genes: 3 *cry8* copies previously reported, 2 *vip1,* and 2 *vip2*. Also, it carries a gene encoding a putative atypical Cry protein. These genes are arranged in a region of approximately 105 kbp in size with characteristics of a pathogenicity island with a potential coleopteran-specific insecticide profile. DNA strand composition asymmetry, as determined by GC skew analysis, and the presence of a Rep protein involved in the initiation of replication suggest a bidirectional *theta* mechanism of replication. In addition, many genes involved in conjugation and a CRISPR-Cas system were detected. The pFR260 sequence was deposited in GenBank under accession number KX258624. 

M.F.B. and R.O.Z. directed different aspects of this work but are equally credited authors.

## Introduction

The *Bacillus cereus* group is a genetically related family of Gram-positive, rod-shaped, spore-forming bacteria showing different virulence spectra. Plasmid members of this group have been involved in pathogenicity as they harbor genes encoding different toxins that affect their respective hosts, including mammals and arthropods. This group includes the well-studied species *B. anthracis* (the causative agent of anthrax), the opportunistic foodborne pathogen *B. cereus,* and the enthomopathogen *B. thuringiensis.* These species may have emerged from a common ancestor through genetic rearrangements leading to acquirement of the pathogenic gene sets [González et al., 1981; Hu et al., 2009].

*B. thuringiensis* produces crystal parasporal inclusions composed of insecticidal proteins (Cry/Cyt) in the sporulation phase of its life cycle. In addition, some *B. thuringiensis* strains also produce vegetative insecticidal proteins (Vip) which are secreted during vegetative growth. The study of genomic sequences has revealed that the insecticidal genes of *B. thuringiensis* localize generally in plasmids and can be organized in an arrangement regarded as pathogenicity island (PAI) [He et al., 2011; Murawska



E-Mail karger@karger.com
www.karger.com/mmb

Laura E. Navas
Instituto de Microbiología y Zoología Agrícola, INTA
Nicolás Repetto y De los Reseros s/n
Hurlingham (1686), Buenos Aires (Argentina)
E-Mail navas.laura@inta.gob.ar

**Table 1.** Summary of pFR260 properties

| Property | Value |
|---|---|
| Total size, bp | 260,595 |
| G+C, % | 33.1 |
| CDS, n | 217 |
| Coding density, % | 67 |
| Average gene length, bp | 813 |
| Transposition-related genes, n | 25 |
| Hypothetical proteins, n | 23 |
| Insecticidal genes, n | 7 |

et al., 2013]. Crystals and soluble secreted proteins have a high specificity against insect hosts and they have gained importance worldwide as an environmentally safe alternative to chemical insecticides to control insect pests [Sauka and Benintende, 2008]. The usefulness of these proteins has also prompted the search for new *B. thuringiensis* isolates in order to identify and characterize insecticidal proteins with new specificities. It is also important to find alternatives to overcome the emergence of insect genetic resistance, particularly in transgenic crops expressing *B. thuringiensis* genes. In recent years, due to the introduction of revolutionary next-generation sequencing technology, a large number of genomes from *B. thuringiensis* isolates have been sequenced [Doggett et al., 2013; He et al., 2011; Murawska et al., 2013] in order to discover new insecticidal proteins useful for biocontrol of agricultural pests and vectors of diseases.

*B. thuringiensis* INTA Fr7-4 is a strain isolated from a soil sample from the province of Misiones, Argentina. It produces ovoid crystals with different levels of toxicity to Lepidoptera and Coleoptera [Amadio et al., 2013; Navas et al., 2014]. We previously reported the complete sequence of 3 plasmids of this strain, named pFR12, pFR12.5, and pFR55 according to their molecular weight in kilo-base pairs [Amadio et al., 2009]. The 2 smallest plasmids were classified as cryptic as they only encode genes with an obvious function involved in their replication, while pFR55 encodes genes homologous to proteins involved in a conjugative machinery. These plasmids have in common the lack of an insecticidal gene. Nevertheless, we found 3 *cry8* genes in this strain, although their localization in the complete genome remains unknown [Amadio et al., 2013; Navas et al., 2014]. The product of one of these genes, i.e., Cry8Pa3 [Navas et al., 2014], showed significant toxicity against the coleopteran pest of cotton, i.e., *Antonomus grandis.* In this paper, we present the complete sequence and analysis of pFR260, a megaplasmid from *B. thuringiensis* INTA Fr7-4 that encodes the 3 previously described *cry8* genes of this strain. These genes are arranged within a region along with genes encoding other virulence factors known to be toxic mainly to coleopteran insects. This plasmid region exhibits features of a PAI with a novel combination of insecticidal gene homologs.

## Results and Discussion

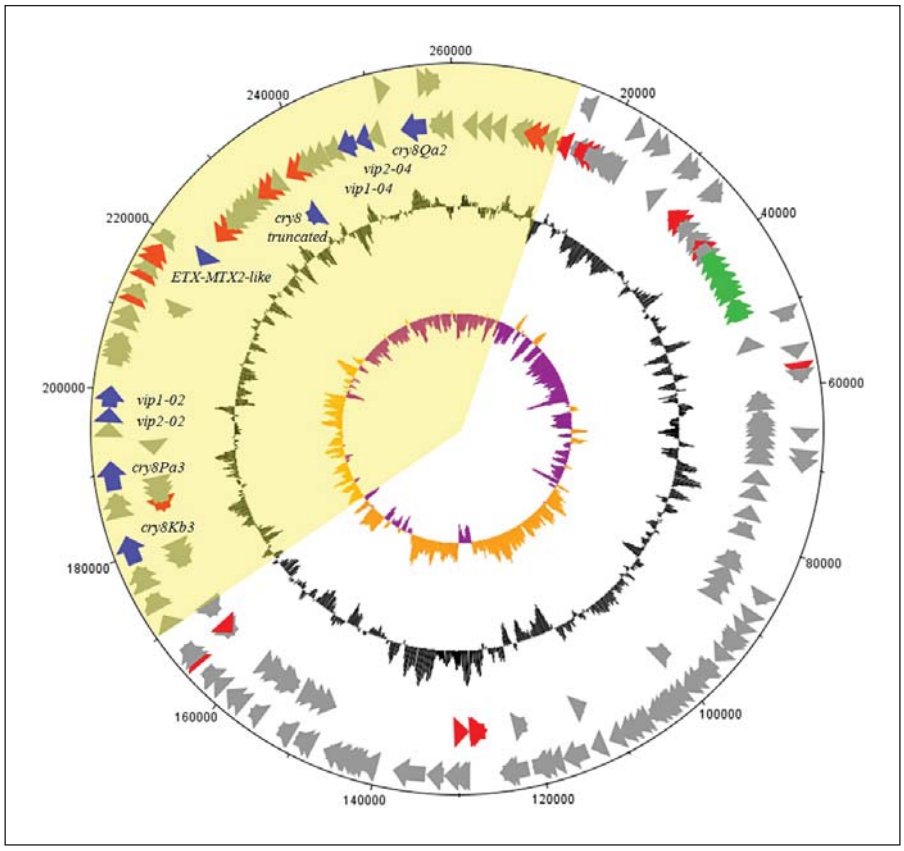
*Overall Features of the pFR260 Megaplasmid*

We have previously identified 3 *cry* genes (i.e., *cry8Qa2*, *cry8Kb3,* and *cry8Pa3*) in *B. thuringiensis* strain INTA Fr7-4 and showed that each gene encodes a protein able to produce crystals when it is expressed individually in a recombinant acrystaliferous strain [Amadio et al., 2013; Navas et al., 2014]. The 3 reported plasmids of this strain do not contain any *cry8* gene. To find the localization of the *cry8* genes, the genome of INTA Fr7-4 was sequenced using Illumina technology.

A de novo genome assembly was done using a Velvet assembler [Zerbino, 2010] and the obtained scaffolds were filtered to select those with sequences similar to those of plasmids from *B. thuringiensis* deposited in GenBank. Four scaffolds that aligned at their ends with the sequences of the *cry8* genes were identified. The implemented gap-closing strategy generated a unique scaffold. The assembly corresponded to the complete sequence of a circular plasmid of 260,595 bp with an overall G+C content of 33.1%. On the basis of its origin and length, the plasmid was named pFR260. Table 1 summarizes the main properties of the plasmid.

A total of 283 coding sequences (CDS) were identified in pFR260 using the RAST v2.0 [Aziz et al., 2008] server, while 250 CDS were annotated using the PROKKA [Seemann, 2014] annotation tool. Each CDS was analyzed individually as described in the Experimental Procedures. As a result, 217 CDS were annotated. Putative functions were assigned to 123 CDS that had significant similarity to proteins of other organisms. Sixty-nine CDS were predicted to encode proteins with similarity to hypothetical proteins, while 25 CDS did not show similarity to reported proteins. tRNA and rRNA were not found in the pFR260 sequence.

Figure 1 shows a representation of pFR260 in which the CDS follow an overall equivalent distribution in both strands (52 and 48%). Only 67% (176,481 bp) of the pFR260 length represents coding regions. This low per-




Color version available online

**Fig. 1.** Circular representation of pFR260. The inner circle represents the GC skew [(G – C)/(G + C)], with positive values in orange and negative ones in purple. The second circle corresponds to the G+C content, with an average value of 33.1% as a base line. The third circle contains only the *cry8*-truncated insecticidal gene. The 2 outer circles represent the predicted genes on both strands. Blue arrows: insecticidal genes; red arrows: transposon-related genes; and green arrows: CRISPR-*cas* genes. The numbers in the outer circle correspond to base pairs. The region highlighted in yellow represents the pathogenicity island. Colors refer to the online version only.

centage of coding density is in agreement with other plasmid studies reporting low gene densities due to the existence of large intergenic regions or truncated genes [Bult et al., 1996; Fraser et al., 1997].

*Plasmid Replication*

pFR260_072 encodes a protein homologous to replicases (Rep) which are essential for *theta*-replicating plasmids [del Solar et al., 1998]. It shows 100% identity to an RepX protein from pBTHD521-5, a plasmid from the *B. thuringiensis* serovar *indiana* HD521 [Li et al., 2015]. The RepX protein of pFR260 possesses the conserved tubulin motif (GGGTGTG) involved in GTP binding that is necessary for protein functionality. Tinsley and Khan [2006] demonstrated that a mutation in this motif abolishes pXO1 replication in *B. anthracis*.

An analysis to predict the directionality and the replication origin of pFR260 was undertaken through a cumulative GC skew [(G – C)/(G + C)] study [Grigoriev, 1998; Necşulea and Lobry, 2007]. DNA strand compositional asymmetry is used to identify bidirectional *theta*-replication origins by detecting a site with a pronounced switch in the cumulative GC-skew, while plasmids replicating through unidirectional *theta* or rolling circle mechanisms have no such switch in their cumulative GC skew [Saillard et al., 2008]. In pFR260, we observed a profile pointing to a bidirectional *theta* mechanism [Arakawa et al., 2009] (online suppl. Fig. 1; for all online suppl. material, see www.karger.com/doi/10.1159/000451056) in which positive values are found along the leader strand and negative ones correspond to the lagged chain. The cumulative GC skew changes sign at the origin and terminus positions. The potential replication origin is located at the coordinate 86,321 (online suppl. Fig. 1), 523 bp away from the 3′ terminus of *repX*. The GC skew can also be observed in Figure 1 (with positive values in orange and negative ones in purple; colors refer to the online version only), where the change in positivity agrees with the mentioned prediction. Furthermore, the region 40 bp upstream and downstream from the putative replication origin is enriched with A+T (~74%), which is characteristic of many replicative regions [Huang et al., 2006]. The high proportion of A+T would facilitate the required strand opening for replication [Rajewska et al., 2012].

**Table 2.** Conjugation-related genes encoded in pFR260

| CDS | Start | End | Strand | Product | Length, amino acids | Organism of origin | Identity, % |
|---|---|---|---|---|---|---|---|
| pFR260_058 | 64,146 | 64,430 | – | membrane protein | 94 | *B. thuringiensis* serovar *morrisoni* | 97 |
| pFR260_059 | 65,261 | 65,524 | – | membrane protein | 87 | *B. thuringiensis* serovar *morrisoni* | 100 |
| pFR260_066 | 70,705 | 72,465 | – | conjugal transfer protein TraG | 586 | *B. cereus* Q1 | 71 |
| pFR260_069 | 78,062 | 78,331 | – | conjugal transfer protein TraG | 89 | *B. cereus* | 98 |
| pFR260_085 | 99,165 | 100,025 | + | Flp pilus assembly protein CpaB | 286 | *B. cereus* | 99 |
| pFR260_088 | 102,303 | 103,235 | + | membrane protein | 310 | *B. thuringiensis* serovar *morrisoni* | 98 |
| pFR260_093 | 104,985 | 105,530 | + | membrane protein | 181 | *B. cereus* AH1134 | 86 |
| pFR260_094 | 105,527 | 106,150 | + | membrane protein | 207 | *B. thuringiensis* serovar *morrisoni* | 98 |
| pFR260_117 | 133,990 | 137,994 | + | membrane protein | 1,334 | *B. thuringiensis* serovar *indiana* | 98 |
| pFR260_121 | 143,477 | 143,860 | + | membrane protein | 127 | *B. cereus* | 98 |
| pFR260_127 | 147,108 | 148,253 | + | membrane protein | 381 | *B. cereus* | 97 |

Most plasmids of the *B. cereus* group belong to the pAMβ1 family of replicons, which follow unidirectional *theta* replication [Bruand et al., 1993; Huang et al., 2006; Wilcks et al., 1999]. There is a lack of information in the available literature about plasmids from *B. thuringiensis* replicating via a bidirectional *theta* mechanism.

*Conjugation-Related Genes*

In analyses of the mechanism of conjugation in plasmids of the *B. cereus* group, the genes responsible for the DNA transfer are grouped in a plasmid region called *tra* [Poluektova et al., 2008; van der Auwera et al., 2005]. This region encodes proteins of the type IV secretion system (T4SS) responsible of DNA transportation between cells.

In plasmid pFR260 we found some putative conjugation-related genes, which are listed in Table 2. We identified the TraG conjugal protein homolog of VirD4. In *Agrobacterium tumefaciens,* VirD4 is involved in energy coupling for DNA transfer in the plasmid conjugative machinery [van der Auwera et al., 2005]. pFR260 also encodes a pilus assembly protein (CpaB) and 8 membrane proteins that could be implicated in cell-cell contact.

Plasmids with functional conjugation systems encode other proteins that take part in T4SS as VirD2, VirB1, VirV11, and VirB4. Homologs of these genes were not found in pFR260, indicating that this plasmid does not code for all proteins required for a complete conjugative cell machine. However, we have reported a plasmid (pFR55) in *B. thuringiensis* INTA Fr7-4 with a putative complete conjugative machinery and described the genes involved in the *tra* region [Amadio et al., 2009]. We speculate that the presence of plasmid pFR55 provides the proteins needed to complete a conjugative cell machinery for pFR260 in *B. thuringiensis* INTA Fr7-4. pFR260 could have lost those genes from its sequence without losing its conjugating ability. Further experiments are required to demonstrate the conjugation ability for the plasmids of *B. thuringiensis* INTA Fr7-4.

*CRISPR-Cas System*

CRISPR (clustered regularly interspaced short palindromic repeats) are found in approximately 45% of sequenced bacterial genomes and 84% of sequenced archaeal genomes (http://crispr.u-psud.fr/crispr/CRISPR-database.php). The CRISPR-Cas system provides acquired immunity against foreign genetic elements such as plasmids and phages [Barrangou et al., 2007; Marraffini and Sontheimer, 2008]. CRISPR are often adjacent to a *cas* (CRISPR-associated) operon that encodes proteins involved in acquisition of that foreign DNA into the CRISPR loci and in the interference stage in an analogous way to iRNA in eukaryotic organisms [Marraffini and Sontheimer, 2010].

In pFR260 we identified a CRISPR/Cas system that consists of 3 CRISPR loci and 9 *cas* genes, 7 of which are comprised of an operon (online suppl. Fig. 2A). Each CRISPR locus contains direct repeats separated by short segments of spacer sequences of exogenous origin. For example, in pFR260, CRISPR 2 contains 10 direct repeats and 9 spacers 33 bp in length (online suppl. Fig. 2B).

There are 3 types of CRISPR/Cas systems depending on the mechanism of interference of exogenous DNA and the different *cas* genes found [Makarova et al., 2011]. Table 3 shows the properties of *cas* genes present in pFR260 and their assigned functions. The CRISPR sequences found in pFR260 belong to the type I subtype C system

**Table 3.** CRISPR-associated genes encoded in pFR260 and their features

| CDS | Start | End | Strand | Product | Length, amino acids | Organism of origin | Identity, % | Function [Horvath and Barrangou, 2010] |
|---|---|---|---|---|---|---|---|---|
| pFR260_039 | 39,736 | 39,960 | – | Csd2 (Cas7) | 74 | *B. cereus* | 98 | association with Cas5, Cas8, and mature |
| pFR260_040 | 40,741 | 41,079 | – | Csd2 (Cas7) | 112 | *B. cereus* F65185 | 98 | RNA for interference with the invading DNA |
| pFR260_041 | 42,038 | 42,328 | – | Cas2 | 96 | *B. cereus* | 100 | recognition of invading DNA and integrati- |
| pFR260_042 | 42,338 | 43,369 | – | Cas1 | 343 | *B. cereus* | 99 | on into a CRISPR locus |
| pFR260_043 | 43,366 | 44,025 | – | Cas4 | 219 | *B. cereus* VD169 | 98 | not determined |
| pFR260_044 | 44,015 | 44,875 | – | Csd2 (Cas7) | 286 | *B. cereus* | 99 | association with Cas5 and mature RNA for |
| pFR260_045 | 44,878 | 46,797 | – | Cas8 | 639 | *B. cereus* VD156 | 99 | interference with the invading DNA |
| pFR260_046 | 46,798 | 47,517 | – | Cas5 | 239 | *B. cereus* | 100 | RNA endonuclease activity for processing of the CRISPR transcript into a mature RNA |
| pFR260_047 | 47,686 | 50,115 | – | Cas3 | 809 | *B. cereus* | 99 | helicase/endonuclease, dsDNA degradation |

due to the combination of the distinct subtype-specific *cas* genes. A similar CRISPR/Cas subtype I-C system has been described for *B. halodurans* [Nam et al., 2012a, b]. Sequences from CRISPR/Cas systems from *B. thuringiensis* are publicly available in databases, although none of them has been studied functionally.

Comparison of the CRISPR/Cas system identified in pFR260 with that of other *B. thuringiensis* plasmids resulted in the identification of 2 plasmids with a similar CRISPR/Cas subtype I-C system. The most similar arrangement was detected in pBT1850294 from the *B. thuringiensis* strain Bt185, the DNA sequences of which show 97.6% identity on average with the 7 *cas* genes of the pFR260 operon [Shu et al., 2009]. Considering insecticidal genes, pBT1850294 also carries the *cry8* genes *cry8Ea1* and *cry8Fa1*. A similar *cas* gene arrangement was detected in pBTHD521-5 from *B. thuringiensis* serovar *indiana* strain HD521 [Li et al., 2015] with 5 genes homologous to those present in pFR260.

*Insecticidal Genes*

pFR260 harbors 7 insecticidal genes with sequence similarity to toxins within 3 different classes, i.e., 3 *cry8* genes, 2 *vip1* genes, and 2 *vip2* genes. pFR260_163, pFR260_170, and pFR260_212 encode the respective proteins Cry8Kb3, Cry8Pa3, and Cry8Qa2, which have been characterized previously [Amadio et al., 2013; Navas et al., 2014].

Along with these genes, pFR260 also contains an ORF encoding 182 amino acids with 50% identity to a fragment covering part of domains 2 and 3 of toxin Cry8Ba (accession No. WP_003305800). The finding of an apparent remnant of a *cry8* gene suggests the loss of part of the gene during the evolution of pFR260. The present composition of the insecticidal genes could represent a dynamic state during the formation of new insecticidal genes within the great diversity in toxins observed in *B. thuringiensis* isolates. The fact that this remnant is located close to CDS with possible roles in transposition implies that transposition is the most probable mechanism to explain the insecticidal gene distribution in different strains [De Maagd et al., 2003], which is consistent with previous observations of *B. thuringiensis* insecticidal genes flanked by transposase sequences [Mahillon et al., 1994; Palma et al., 2014a]. Plasmid pBtoxis from *B. thuringiensis* serovar israelensis [Berry et al., 2002] also presents insecticidal gene remnants in correlation with a high number of transposition-related genes. Overall, 12% of the complete CDS of pFR260 and 57% of truncated genes presented similarities to transposition-related genes, suggesting that a considerable amount of DNA transposition has occurred during the evolution of pFR260.

In addition to *cry8* genes, pFR260_188 encodes a protein with an ETX_MTX2 conserved domain. This domain is present in some Cry toxins (Cry15, Cry23, Cry33, Cry38, Cry45, Cry51, Cry60, and Cry64) from *B. thuringiensis* homologous to *Lysinibacillus sphaericus* toxins [Berry, 2012] and to the epsilon toxin of *Clostridium* [Bokori-Brown et al., 2011]. Proteins of this group are predicted to fold into 1 domain tertiary structure in contrast to the 3 domain structures of most Cry proteins. However, the mechanism of action seems to be similar, including pore formation [Adang et al., 2014]. The deduced amino acid sequence of pFR260_188 has a predicted molecular weight of 36 kDa and displays its highest identity (35%) with the mosquitocidal toxin Mtx2/3 from *L.*

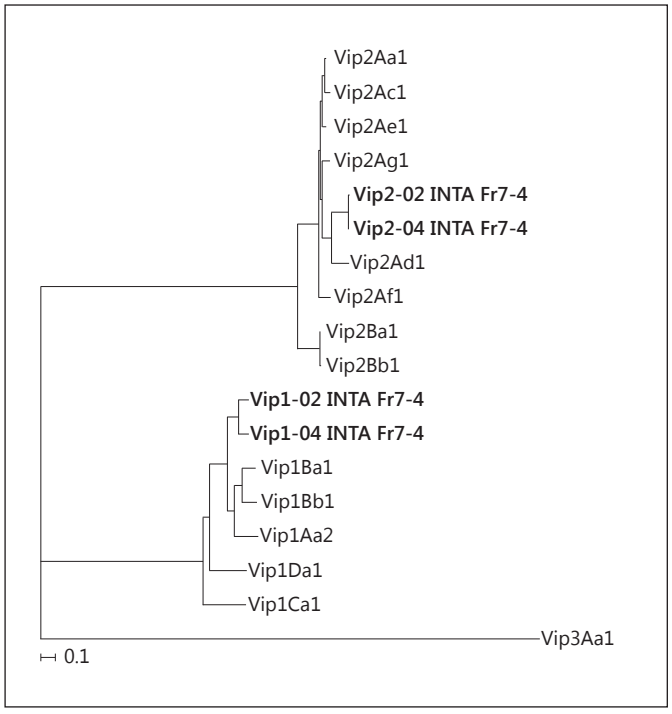



**Fig. 2.** Neighbor-joining tree showing the phylogenetic relationship among different available Vip1 and Vip2 proteins. Vip3Aa1 was included as the out group.

*sphaericus* [Berry, 2012]. An intraspecies comparison to other Cry toxins featuring the ETX_MTX2 domain in *B. thuringiensis* showed identity levels below 10%.

pFR260_173/pFR260_174 and pFR260_208/pFR260_209 are homologous to the binary Vip, i.e., *vip2/vip1* pairs, produced by *B. thuringiensis* isolates that are secreted during the vegetative growth phase. They represent 2 different *vip2/vip1* operons according to their arrangement in tandem, similarly to genes of this group reported previously [Chakroun et al., 2016]. It has been proposed that the cytotoxic protein corresponds to Vip2, while Vip1 would be involved in receptor binding and transportation of Vip2 to the cytoplasm of the cell [De Maagd et al., 2003]. The predicted molecular weights are according to those described for Vip1 and Vip2 proteins [Palma et al., 2014b].

All Vip proteins from INTA Fr7-4 share an identity lower than 68% with reported sequences. Vip1 proteins from *B. thuringiensis* INTA Fr7-4 were named Vip1-02 and Vip1-04 and have a predicted molecular weight of around 97 kDa. They share an overall 85% sequence identity, with differences located mainly in the last 250 amino acids of the C-terminus, as observed by Shi et al. [2004] for other Vip1 proteins. Vip2 proteins from the INTA Fr7-4 strain, named Vip2-02 and Vip2-04, possess a predicted molecular weight of 55 kDa. They share an overall 99% sequence identity, and the most variable region with the reported Vip2 proteins is located in a 60- to 70-amino acid region from the N-terminus.

Vip2Ad1 is the most similar protein to both Vip2 from INTA Fr7-4 (Fig. 2), with a sequence identity of 74 and 73% with Vip2-04 and Vip2-02, respectively. It has been isolated from *Brevibacillus laterosporus* [Boets et al., 2004] with reported toxicity against coleopteran insects.

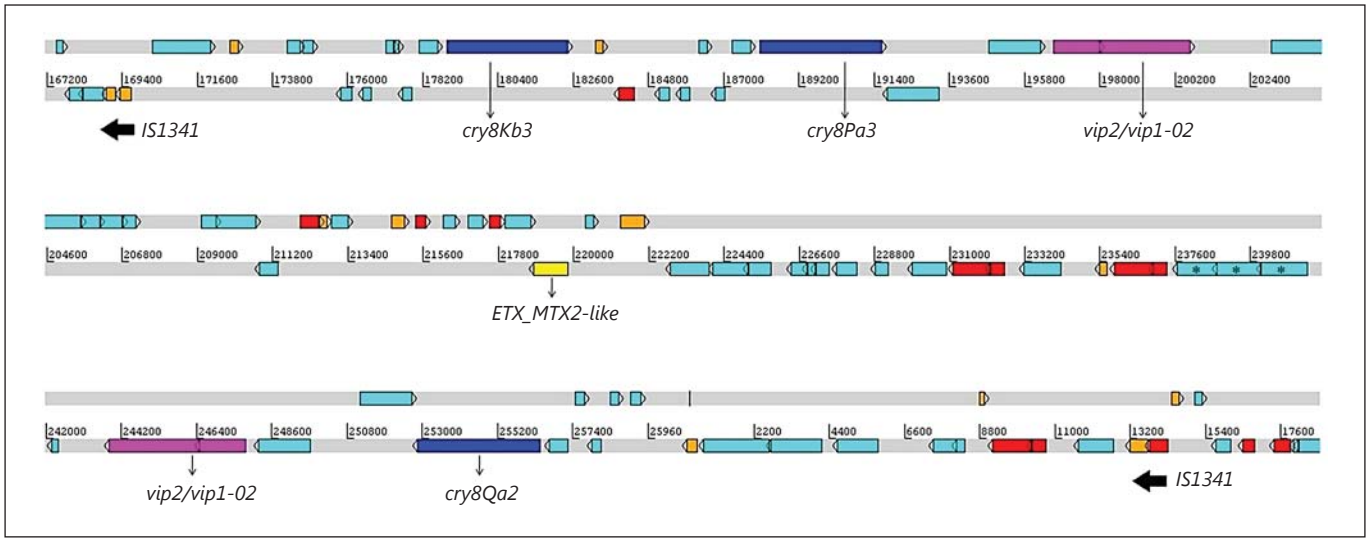
Vip1Bb1 from the *B. thuringiensis* strain PS177C8 is the most similar protein to Vip1-02 reported from other isolates, with 68% identity, whereas Vip1Aa2 from the *B. thuringiensis* strain PS31F2 is the most similar protein to Vip1-04, with 67% identity. Both proteins are part of a patent reporting 9 *vip* genes toxic to Lepidoptera and Coleoptera [Feitelson et al., 1999, 2003].

*Pathogenicity Island*

An important feature of pFR260 is the distribution of the insecticidal genes within a ~105-kbp region that exhibits most of the characteristics of a bacterial PAI as defined by Hacker and Kaper [2000]. This includes: the presence of all of the virulence factors of the strain; a high content of genes or pseudo-genes related to mobility such as integrases, transposases, and IS; the boundaries of the PAI composed by directed repeats; the presence of proteins related to secretion systems; and a total length of around 10–200 kbp [Hacker and Kaper, 2000].

Focusing on the insecticidal genes within the PAI, we reported previously that *cry8Kb3* and *cry8Pa3* are coded in the same strand separated by ~5.6 kbp [Navas et al., 2014]. One of the *vip2/vip1* operons reported in this work is located ~5 kbp downstream from the *cry8Pa3* gene. Therefore, *cry8Kb3, cry8Pa3, vip2-02,* and *vip1-02* are all coded in proximate distances on the same strand (Fig. 3). On the opposite strand, ~44 kbp downstream, the *vip2/vip1-04* operon and the previously reported *cry8Qa2* [Amadio et al., 2013] are located in a relative position similar to that described for their homologs in the other strand. The regions between *cry8* genes and *vip* operons are highly conserved. Between both *vip* operons pFR260_188, the potential ETX_MTX2 domain toxin, is coded. The region corresponding to the PAI in pFR260 is highlighted in Figure 1 in addition to Figure 3.

Cry8 proteins with toxicity against coleopteran insects, including Cry8Pa3 from *B. thuringiensis* INTA Fr7-4 [Navas et al., 2014], have been reported [e.g., Asano et




Color version available online

**Fig. 3.** pFR260 physical map of the PAI. The arrows indicate the predicted CDS. Blue arrows: *cry8* genes; purple arrows: *vip* genes; yellow arrow ETX_MTX2-like gene; red arrows: transposition-related genes; orange arrows: transposition-related truncated genes; light blue arrows: CDS with another function; black arrows: region with similarity to IS1341 that indicates the probable PAI boundaries. The asterisks indicate the hemolysin genes. Coordinates are represented between both strands. Colors refer to the online version only.

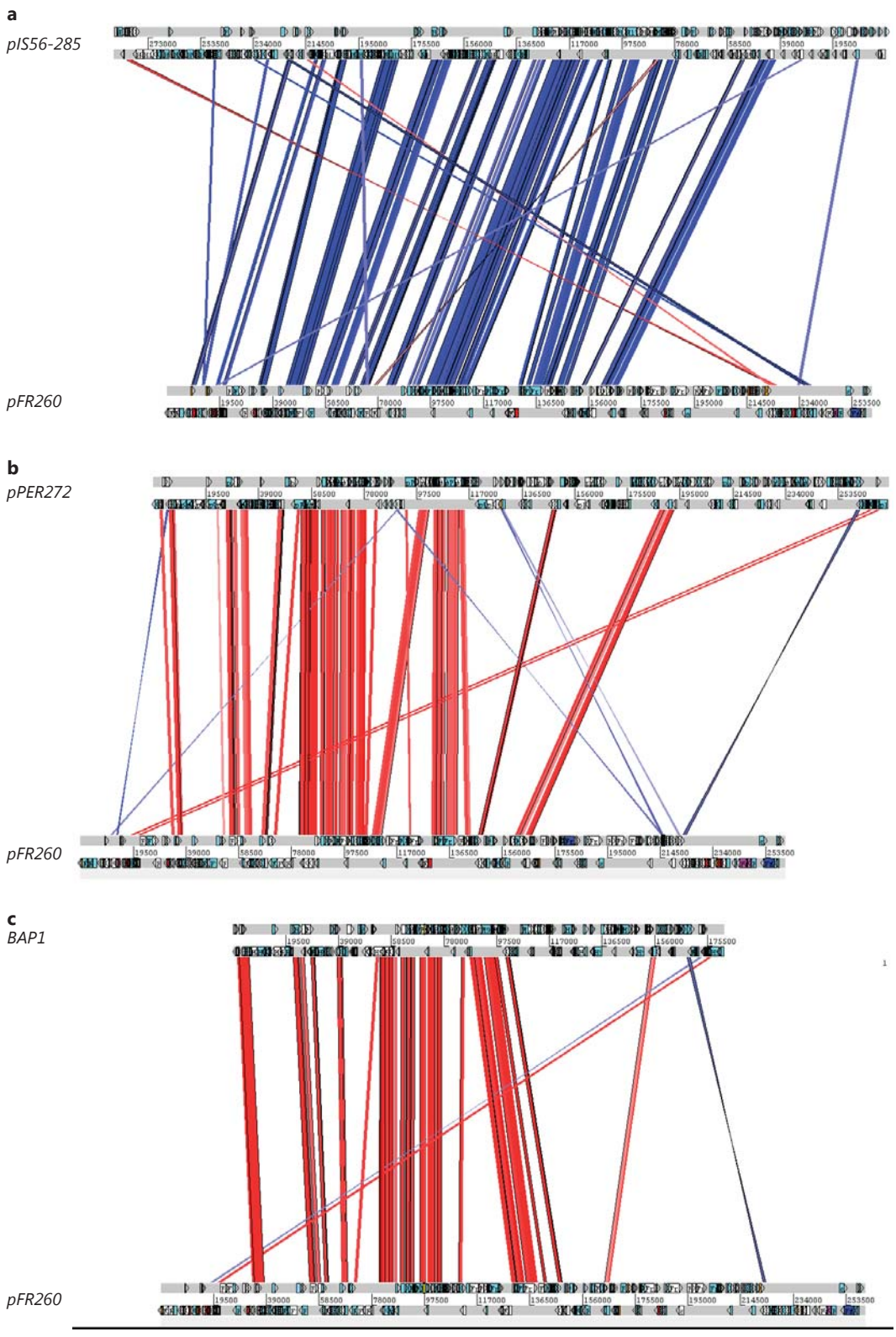
al., 2003; Oliveira et al., 2011; Shu et al., 2009]. The binary toxins Vip2/Vip1 have been found to have specificity to Coleoptera [Arnaut et al., 2011; Bi et al., 2015; Feitelson et al., 2003; Schnepf et al., 2003; Shingote et al., 2013; Warren, 1997]. These families of genes have been observed to occur together [Hernández-Rodríguez et al., 2009]. In particular, Bi et al. [2015] described a *B. thuringiensis* strain toxic to *Scarabaeidae* larvae carrying a *cry8* gene, a *cry8-like* gene, and a *vip1*/*vip2* operon, but no description about the organization of these genes in the genome was given. On the other hand, most of the PAI described for *B. thuringiensis* [He et al., 2011; Murawska et al., 2013] include *cry1A, cry1I, cry2A,* and *vip3A* genes, which are typical of Lepidoptera-specific strains. To our knowledge, this is the first report of coleopteran-specific genes *(cry8* and *vip2/vip1)* grouped in a plasmid region with characteristics of a bacterial PAI.

Another virulence factor present in this region of pFR260 is the hemolysin BL, a tripartite well-characterized toxin typically found in *B. cereus* strains [Prüss et al., 1999], although it has been demonstrated to occur in *B. thuringiensis* strains as a pathogenic assistance factor [Zhu et al., 2015]. pFR260_204, pFR260_205, and pFR260_206 are located in the PAI (Fig. 3) and have 97, 99, and 98% identity to the hemolysin BL-binding component, hemolysin BL lytic component L1, and hemolysin BL lytic component L2 from *B. cereus,* respectively. This suggests the ability of *B. thuringiensis* INTA Fr7-4 to produce a different hemolysin, which may contribute to its pathogenicity.

As mentioned before, there are a high number of mobile element-related genes in pFR260. Within this ~105-kbp region there is an enrichment of mobile genetic elements, mainly in the insertion sequences (IS) of the families IS200/IS605, IS3, and IS4. Both limits of this region are flanked by direct repeats that are part of an IS1341, which belong to the IS605 family [He et al., 2015]. A PAI acquired by recombinant events through IS605 elements has been reported for *Helicobacter pylori* [Censini et al., 1996]. For this reason, the presence of IS605 in the boundaries of the pFR260 PAI could be responsible for its being acquired by lateral transfer.

*Comparison of pFR260 with Related Plasmids*

The sequence of pFR260 does not show a global similarity with any plasmid reported to date. In a comparison with databases sequences, partial similarities with plasmids of the *B. cereus* group could be detected. One plasmid of each species (*B. thuringiensis, B. cereus,* and *B. anthracis)* that yielded the best BLAST hit was selected for an individual comparison with pFR260 (Fig. 4a–c).




4

*(For legend see next page.)*

pIS56-285 from the *B. thuringiensis* serovar *thuringiensis* strain IS5056 shows detectable similarity with 135 of the 217 predicted proteins in pFR260 (Fig. 4a). The host strain was isolated from soil in Biebrza National Park (Poland) and produces a quasicuboidal bipyramidal crystal composed of Cry1Ab21 highly toxic to *Trichoplusia ni* larvae. Similarly to pFR260, pIS56-285 carries a PAI that differs in the class and number of insecticidal genes, consisting of 2 *cry1,* 2 *cry2,* and a *vip3* gene, giving a lepidopteran-specific profile [Murawska et al., 2013]. The region of ~105 kbp corresponding to the PAI in pFR260 is less conserved than the rest of the plasmid (Fig. 4a). The few alignments within this region correspond to transposase genes and the tripartite hemolysin. The limits of both plasmid PAIs are conserved, encoding transposases. This comparison shows that the plasmid backbone is moderately conserved while the PAIs are the most different regions between them. The differences in the insecticidal gene families found in pIS56-285 and pFR260 are in agreement with the different specificity against different orders of insects.

Figure 4b and c shows alignment between pFR260 and the most similar plasmids from *B. cereus* and *B. anthracis,* respectively. They are: pPER272 from the *B. cereus* strain AH820 and BAP1 from the *B. anthracis* strain H9401. Overall, 75 and 83 of the 217 predicted pFR260 proteins showed detectable similarity to those plasmids, respectively, encoding mainly membrane proteins, transposases, and hypothetical and conjugation-related proteins.

The comparison analysis demonstrated that pFR260 is a novel plasmid with similarities to other plasmids of the *B. cereus* group and differences mainly in the PAI region with potential coleopteran-specific insecticidal genes.

In conclusion, *B. thuringiensis* INTA Fr7-4 harbors at least 4 plasmids – 3 of them are small plasmids previously reported and the fourth is a megaplasmid of 260,595 bp, i.e., pFR260, described in this study. It contains 217 CDS, 40% of which have an unknown function and 12% of which are related to mobile elements. pFR260 is characterized by the presence of a ~105-kbp region with features of a PAI, containing 3 *cry8* genes, 2 *vip1,* and 2 *vip2,* different from each other, and a gene encoding an ETX_MTX2 domain protein that may represent a new insecticidal protein family. This gene profile attributes a putative coleopteran-specific activity to the strain, and it is in agreement with the toxicity against *Anthonomus grandis* previously detected [Navas et al., 2014]. pFR260 represents a novel megaplasmid from *B. thuringiensis* species.

**Fig. 4.** Comparison of pFR260 with related plasmids. Linear representation of the pFR260 alignment with pIS56-285 (**a**; accession No. NC_020384.1), pPER272 (**b**; accession No. DQ889678.1), and BAP1 (**c**; accession No. CP002092.1). Protein-protein similarities (as determined by TBLASTX comparisons of the complete plasmids) are indicated by lines, with the strength of the match indicated by the intensity of the color. Lines in blue indicate that a match is made with the reverse complement of the sequence. Color refers to the online version only.

## Experimental Procedures

*Strain and DNA Preparation*

*B. thuringiensis* INTA Fr7-4 was obtained from the IMyZA-INTA bacterial collection. Total genomic DNA was purified as follows: a 100-mL overnight culture of *B. thuringiensis* INTA Fr7-4 in LB medium was centrifuged at 4.080 *g* for 15 min at 4°C and resuspended in 10 mL TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8) containing 10 mg/mL lysozyme. After 20 min of incubation at room temperature, SDS and RNAse A were added to a final concentration of 1% and 40 µg/mL, respectively. After incubation for 1 h at 37°C, 5 mL 5 M NaCl was added and mixed thoroughly. Three milliliters of 10% CTAB (cetyl trimethylammonium bromide) in 0.7 M NaCl was added, followed by incubation for 20 min at 65°C. DNA was extracted twice with an equal volume of chloroform/isoamyl alcohol (24:1). DNA was precipitated with 0.6 volumes of isopropanol. A stringy white DNA precipitate was transferred to a tube containing 70% ethanol using a disposable inoculating loop. This wash was repeated twice. The dried pellet was dissolved in TE buffer.

*Sequencing, Assembly, and Gap Closure*

Total DNA from *B. thuringiensis* INTA Fr7-4 was sequenced using a long jumping distance library of 8 kbp and 2 × 150 bp run on a MiSeq Illumina apparatus. De novo assembly was performed using Velvet [Zerbino, 2010]. The obtained scaffolds were analyzed using BLASTN to filter only those with similarities to plasmids deposited in the GenBank database. Within this set, we identified scaffolds aligned to the previously reported *cry8* genes [Amadio et al., 2013; Navas et al., 2014]. The sequences of these genes were used to guide the assembly of selected scaffolds and gap closure resulting in a unique scaffold representing a plasmid. In silico gap filling was performed with GapFiller 1.10 [Boetzer and Pirovano, 2012], and the remaining gaps were closed using a primer walking strategy. Primers were designed using a function included in the GAP4 program from the Staden Package [Staden et al., 1999]. AccuPrime™ High-Fidelity Taq DNA Polymerase (Invitrogen) was used for PCR following the manufacturer's specifications. The annealing temperature was calculated using FastPCR software according to the thermodynamic properties of the primers [Kalendar et al., 2009]. PCR products were sequenced using an ABI3130xl (Applied Biosystem) sequencer. Sequences obtained by the Sanger chain termination method were assembled onto the scaffold using GAP4.

*Sequence Annotation*

Automated annotation was performed using PROKKA [Seemann, 2014], which uses Prodigal [Hyatt, 2010] as a gene-finder tool and BLAST [Altschul et al., 1997] to make an alignment when comparing to a database. We also tested the RAST server [Aziz et al., 2008] annotation tool, which uses Glimmer3 [Delcher et al., 1999] for gene prediction and makes the function assignment in comparison to FIGfams families. The results were integrated and corrected manually in Artemis v16.0 [Rutherford et al., 2000]. Manual curation included the analysis of BLASTP hits for each CDS against the GenBank, Uniprot, and PFam databases. Function was assigned if the CDS and the hit shared at least 30% amino acid identity across at least 70% of the CDS length. CDS displaying less identity were annotated as "hypothetical proteins", whereas those displaying sequence identity to a gene of unknown function were annotated as "conserved hypothetical proteins". The truncated genes were defined on the basis of the criterion for function assignment when the hit protein coverage was less than 70%. CDS smaller than 40 amino acids and truncated CDS were discarded.

*Toxin Genes Analysis*

Toxin genes were identified using the BtToxin_Scanner tool (http://bcam.hzaubmb.org/BtToxin_scanner). Multiple sequence alignment was generated using ClustalX2 [Larkin et al., 2007]. The model of amino acid change used to compute the distance matrix was that of Jones et al. [1992], and the neighbor-joining method of clustering of Saitou and Nei [1987] was used to construct the phylogenetic tree. TreeView was used to plot the rooted tree diagram [Page, 1996]. The Vip3Aa1 sequence (GenBank accession No. AAC37036) was used as the out group.

*Plasmid Analysis*

A circular diagram of pFR260 was created using DNAPlotter [Carver et al., 2009]. The cumulative GC skew graphic used for prediction of the origin of replication was plotted using the GenSkew tool (http://genskew.csb.univie.ac.at). Comparisons among related plasmids were made with BLAST [Altschul et al., 1997] and visualized using the Artemis Comparison Tool [Carver et al., 2005].

## Nucleotide Sequence Accession Number

The circularized sequence was deposited in GenBank under the accession number KX258624.

## Acknowledgments

We thank Irma Fuxan for the technical support in sequencing and routine laboratory protocols. L.E.N. and E.M.O. were supported by a fellowship from the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). A.F.A., D.H.S., M.F.B., and R.O.Z. are staff members of CONICET. This work was funded by project FONCYT PICT 2010-0356.

## References

Adang MJ, Crickmore N, Jurat-Fuentes JL: Diversity of *Bacillus thuringiensis* crystal toxins and mechanism of action; in Dhadialla T, Gill S (eds): Insect Midgut and Insecticidal Proteins: Advances in Insect Physiology. London, Academic Press, 2014, vol 47, pp 39–87.

Altschul SF, Madden TL, Schäffer AA, Zhang J, et al: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997;25:3389–3402.

Amadio AF, Benintende GB, Zandomeni RO: Complete sequence of three plasmids from *Bacillus thuringiensis* INTA-FR7-4 environmental isolate and comparison with related plasmids from the *Bacillus cereus* group. Plasmid 2009;62:172–182.

Amadio AF, Navas LE, Sauka DH, Berretta MF, Benintende GB, Zandomeni RO: Identification, cloning and expression of an insecticide *cry8* gene from *Bacillus thuringiensis* INTA Fr7-4. J Mol Microbiol Biotechnol 2013;23:401–409.

Arakawa K, Suzuki H, Tomita M: Quantitative analysis of replication-related mutation and selection pressures in bacterial chromosomes and plasmids using generalised GC skew index. BMC Genomics 2009;10:640.

Arnaut G, Boets A, Damme N, Van Rie J: Toxins. US Patent No 7,919,609. 2011.

Asano S, Yamashita C, Iizuka T, Takeuchi K, Yamanaka S, Cerf D, Yamamoto T: A strain of *Bacillus thuringiensis* subsp. *galleriae* containing a novel *cry8* gene highly toxic to *Anomala cuprea* (Coleoptera: Scarabaeidae). Biol Control 2003;28:191–196.

Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O: The RAST server: rapid annotations using subsystems technology. BMC Genomics 2008;9:1–75.

Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P: CRISPR provides acquired resistance against viruses in prokaryotes. Science 2007;315:1709–1712.

Berry C, O'Neil S, Ben-Dov E, Jones AF, et al: Complete sequence and organization of pBtoxis, the toxin-coding plasmid of *Bacillus thuringiensis* subsp. *israelensis*. Appl Environ Microbiol 2002;68:5082–5095.

Berry C: The bacterium, *Lysinibacillus sphaericus,* as an insect pathogen. J Invertebr Pathol 2012;109:1–10.

Bi Y, Zhang Y, Shu C, Crickmore N, Wang Q, Du L, Song F, Zhang J: Genomic sequencing identifies novel *Bacillus thuringiensis* Vip1/Vip2 binary and Cry8 toxins that have high toxicity to Scarabaeoidea larvae. Appl Microbiol Biotechnol 2015;99:753–760.

Boets A, Arnaut G, Van Rie J, Damme N: Toxins. US Patent No 6,706,860. 2004.

Boetzer M, Pirovano W: Toward almost closed genomes with GapFiller. Genome Biol 2012;13:R56.

Bokori-Brown M, Savva CG, Fernandes da Costa SP, Naylor CE, Basak AK, Titball RW: Molecular basis of toxicity of *Clostridium perfringens* epsilon toxin. FEBS J 2011;278:4589–4601.

Bruand C, Lechatelier E, Ehrlich SD, Jannière L: A fourth class of theta-replicating plasmids: the pAM$\beta_1$ family from Gram-positive bacteria. Proc Natl Acad Sci USA 1993;90:11668–11672.

Bult C, White O, Olsen GJ, Zhou L, Fleischmann RD, Sutton GG, Blake JA, FitzGerald LM, Clayton RA, Gocayne JD, Kerlavage AR, Dougherty BA, Tomb JF, Adams MD, Reich CI, Overbeek R, Kirkness EF, Weinstock KG, Merrick JM, Glodek A, Scott JL, Geoghagen NS. Venter JC: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii.* Science 1996;273:1058–1073.

Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, Parkhill J: ACT: the Artemis Comparison Tool. Bioinformatics 2005;21:3422–3423.

Carver TJ, Thomson N, Bleasby A, Berriman M, Parkhill J: DNAPlotter: circular and linear interactive genome visualization. Bioinformatics 2009;25:119–120.

Censini SC, Lange Z, Xiang JE, Crabtree P, Ghiara M, Borodovsky R, Rappuoli, Covacci A: cag, a pathogenicity island of *Helicobacter pylori*, encodes type I- specific and disease-associated virulence factors. Proc Natl Acad Sci USA 1996;93:14648–14653.

Chakroun M, Banyuls N, Bel Y, Escriche B, Ferré J: Bacterial vegetative insecticidal proteins (Vip) from entomopathogenic bacteria. Microbiol Mol Biol Rev 2016;80:329–350.

De Maagd RA, Bravo A, Berry C, Crickmore N, Schnepf HE: Structure, diversity, and evolution of protein toxins from spore-forming entomopathogenic bacteria. Annu Rev Genet 2003;37:409–433.

Delcher AL, Harmon D, Kasif S, White O, Salzberg SL: Improved microbial gene identification with GLIMMER. Nucleic Acids Res 1999; 27:4636–4641.

del Solar G, Giraldo R, Ruiz-Echevarria MJ, Espinosa M, Diaz-Orejas R: Replication and control of circular bacterial plasmids. Microbiol Mol Biol Rev 1998;62:434–464.

Doggett NA, Stubben CJ, Chertkov O, Bruce DC, Detter JC, Johnson SL, Han CS: Complete genome sequence of *Bacillus thuringiensis* serovar *israelensis* strain HD-789. Genome Announc 2013, DOI: 10.1128/genomeA.01023-13.

Feitelson JS, Schnepf HE, Narva KE, Stockhoff BA, et al: Plants and cells transformed with a nucleic acid from *Bacillus thuringiensis* strain KB59A4-6 encoding a novel SUP toxin. US Patent No 6,603,063. 1999.

Feitelson JS, Schnepf HE, Narva KE, Stockhoff BA, Schmeits J, Loewer D, Dullum CJ, Muller-Cohn J, Stamp L, Morrill G, Finstad-Lee S: Pesticidal toxins and nucleotide sequences which encode these toxins. US Patent No 6,656,908. 2003.

Fraser CM, Casjens S, HuangWM, Sutton GG, Clayton R, Lathigra R, White O, Ketchum KA, Dodson R, Hickeyt EK, Gwinn M, Dougherty B, Tomb JF, Fleischmann RD, Richardson D, Peterson J, Kerlavage AR, Quackenbush J, Salzberg S, Hanson M, van Vugt R, Palmer N, Adams MD, Gocayne J, Weidman J, Utterback T, Watthey L, McDonald L, Artiach P, Bowman C, Garland S, Fuji C, Cotton MD, Horst K, Roberts K, Hatch B, Smith HO, Venter JC: Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*. Nature 1997;390:580–588.

González J, Dulmage H, Carlton B: Correlation between specific plasmids and delta-endotoxin production in *Bacillus thuringiensis*. Plasmid 1981;5:352–365.

Grigoriev A: Analyzing genomes with cumulative skew diagrams. Nucleic Acids Res 1998;26: 2286–2290.

Hacker J, Kaper JB: Pathogenicity islands and the evolution of microbes. Annu Rev Microbiol 2000;54:641–679.

He J, Wang J, Yin W, Shao X, Zheng H, Li M, Zhao Y, Sun M, Wang S, Yu Z: Complete Genome Sequence of *Bacillus thuringiensis* subsp. *chinensis* strain CT-43. J Bacteriol 2011;193: 3407–3408.

He S, Corneloup A, Guynet C, Lavatine L, Caumont-Sarcos A, Siguier P, Marty B, Dyda F, Chandler M, Ton Hoang B: The IS200/IS605 family and 'peel and paste' single-strand transposition mechanism. Microbiol Spectr 2015, DOI: 10.1128/microbiolspec.MDNA3-0039-2014.

Hernández-Rodríguez CS, Boets A, Van Rie J, Ferré J: Screening and identification of *vip* genes in *Bacillus thuringiensis* strains. J Appl Microbiol 2009;107:219–225.

Horvath P, Barrangou R: CRISPR/Cas, the immune system of bacteria and archaea. Science 2010;327:167–170.

Hu X, Swiecicka I, Timmery S, Mahillon J: Sympatric soil communities of *Bacillus cereus sensu lato*: population structure and potential plasmid dynamics of pXO1- and pXO2-like elements. FEMS Microbiol Ecol 2009;70:344–355.

Huang J, Guo S, Mahillon J, Van der Auwera GA, et al: Molecular characterization of a DNA fragment harboring the replicon of pBMB165 from *Bacillus thuringiensis* subsp. *tenebrionis*. BMC Genomics 2006;7:270.

Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ: Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 2010; 11:119.

Jones DT, Taylor WR, Thornton JM: The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 1992; 8:275–282.

Kalendar R, Lee D, Schulman AH: FastPCR software for PCR primer and probe design and repeat search. Genes Genomes Genomics 2009;3:1–14.

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG: Clustal W and Clustal X version 2.0. Bioinformatics 2007;23: 2947–2948.

Li Q, Xu LZ, Zou T, Ai P, Huang GH, Li P, Zheng AP: Complete genome sequence of *Bacillus thuringiensis* strain HD521. Stand Genomic Sci 2015;10:62.

Mahillon JR, Rezsöhazy B, Hallet B, Delcour J: IS231 and other *Bacillus thuringiensis* transposable elements: a review. Genetica 1994;93: 13–26.

Makarova KS, Haft DH, Barrangou R, Brouns SJ, Carpentier E, Horvath P, Moineau S, Mojica FJ, Wolf YI, Yakunin AF, van der Oost J, Koonin EV: Evolution and classification of the CRISPR-Cas systems. Nat Rev Microbiol 2011;9:467–477.

Marraffini LA, Sontheimer EJ: CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. Science 2008;322: 1843–1845.

Marraffini LA, Sontheimer EJ: CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. Nat Rev Genet 2010;11: 181–190.

Murawska E, Fiedoruk K, Bideshi DK, Swiecicka I: Complete genome sequence of *Bacillus thuringiensis* subsp. *thuringiensis* strain IS5056, an isolate highly toxic to *Trichoplusia ni*. Genome Announc 2013;1:e0010813.

Nam KH, Haitjema C, Liu X, Din F, et al: Cas5d protein processes pre-crRNA and assembles into a cascade-like interference complex in subtype I-C/Dvulg CRISPR-Cas system. Structure 2012a; 20:1574–1584.

Nam KH, Ding F, Haitjema C, Huang Q, et al: Double-stranded endonuclease activity in *Bacillus halodurans* clustered regularly interspaced short palindromic repeats (CRISPR)-associated Cas2 protein. J Biol Chem 2012b; 287:35943–35952.

Navas LE, Berretta MF, Pérez MP, Amadio AF, Ortiz EM, Sauka DH, Benintende GB, Zandomeni RO: Sequence and expression of two *cry8* genes from *Bacillus thuringiensis* INTA Fr7-4, a native strain from Argentina. J Mol Microbiol Biotechnol 2014;24:241–248.

Necşulea A, Lobry JR: A new method for assessing the effect of replication on DNA base composition asymmetry. Mol Biol Evol 2007;24: 2169–2179.

Oliveira GR, Silva MC, Lucena WA, Nakasu EY, Firmino AA, Beneventi MA, Souza DS, Gomes JE Jr, de Souza JD Jr, Rigden DJ, Ramos. HB, Soccol CR, Grossi-de-Sa MF: Improving Cry8Ka toxin activity towards the cotton boll weevil *(Anthonomus grandis)*. BMC Biotechnol 2011;11:85.

Page RDM: TreeView: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 1996;12:357–358.

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 2014a;6:1490–1504.

Palma L, Muñoz D, Berry C, Murillo J, Caballero P: *Bacillus thuringiensis* toxins: an overview of their biocidal activity. Toxins 2014b; 6:3296–3325.

Poluektova EU, Gagarina E, Shilovskiĭ IP, Fedorina EA, Nezametdinova VZ, Rodionova SA, Prozorov AA: Molecular analysis of some genes from plasmid p19 of the soil strain *Bacillus subtilis* 19 involved in conjugation (in Russian). Genetika 2008;44:623–630.

Prüss BM, Dietrich R, Nibler B, Märtlbauer E, Scherer S: The hemolytic enterotoxin HBL is broadly distributed among species of the *Bacillus cereus* group. Appl Environ Microbiol 1999;65:5436–5442.

Rajewska M, Wegrzyn K, Konieczny I: AT-rich region and repeated sequences – the essential elements of replication origins of bacterial replicons. FEMS Microbiol Rev 2012:36:408–434.

Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B: Artemis: sequence visualization and annotation. Bioinformatics 2000;16:944–945.

Saillard C, Carle P, Duret-Nurbel S, Henri R, Killiny N, Carrère S, Gouzy J, Bové JM, Renaudin J, Foissac X: The abundant extrachromosomal DNA content of the *Spiroplasma citri* GII3-3X genome. BMC Genomics 2008;9-195.

Saitou N, Nei M: The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 1987;4:406–425.

Sauka D, Benintende G: *Bacillus thuringiensis*: general aspects – an approach to its use in the biological control of lepidopteran insects behaving as agricultural pests. Rev Argent Microbiol 2008;40:124–140.

Schnepf HE, Narva KE, Stockhoff BA, Lee SF, Waltz M, Sturgis B: Pesticidal toxins and genes from *Bacillus laterosporus* strains. US Patent No 6,605,701. 2003.

Seemann, T: Prokka: rapid prokaryotic genome annotation. Bioinformatics 2014;30:2068–2069.

Shi Y, Xu W, Yuan M, Tang M, Chen J, Pang Y: Expression of *vip1/vip2* genes in *Escherichia coli* and *Bacillus thuringiensis* and the analysis of their signal peptides. J Appl Microbiol 2004;97:757–765.

Shingote PR, Moharil MP, Dhumale DR, Jadhav PV, Satpute NS, Dudhare MS: Screening of *vip1/vip2* binary toxin gene and its isolation and cloning from local *Bacillus thuringiensis* isolates. Sci Asia 2013;39:620–624.

Shu C, Yu H, Wang R, Fen S, Su X, Huang D, Zhang J, Song F: Characterization of two novel *cry8* genes from *Bacillus thuringiensis* strain BT185. Curr Microbiol 2009;58:389–392.

Staden R, Beal KF, Bonfield JK: The Staden Package, 1998. Methods Mol Biol 1999;132:115–130.

Tinsley E, Khan SA: A novel FtsZ-like protein is involved in replication of the anthrax toxin-encoding pXO1 plasmid in *Bacillus anthracis*. J Bacteriol 2006;188:2829–2835.

van der Auwera GA, Andrup L, Mahillon J: Conjugative plasmid pAW63 brings new insights into the genesis of the *Bacillus anthracis* virulence plasmid pXO2 and of the *Bacillus thuringiensis* plasmid pBT9727. BMC Genomics 2005;6:103.

Warren GW: Vegetative insecticidal proteins: novel proteins for control of corn pests; in Carozzi N, Koziel M (eds): Advances in Insect Control: the Role of Transgenic Plants. London, Taylor and Francis, 1997, pp 109–121.

Wilcks A, Smidt L, Okstad OA, Kolsto AB, Mahillon J, Andrup L: Replication mechanism and sequence analysis of the replicon of pAW63, a conjugative plasmid from *Bacillus thuringiensis*. J Bacteriol 1999;181:3193–3200.

Zerbino DR: Using the Velvet de novo assembler for short-read sequencing technologies. Curr Protoc Bioinformatics. 2010, DOI: 10.1002/0471250953.bi1105s31.

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