Plasmid 62 (2009) 172–182

Contents lists available at ScienceDirect

Plasmid

journal homepage: www.elsevier.com/locate/yplas

Complete sequence of three plasmids from *Bacillus thuringiensis* INTA-FR7-4 environmental isolate and comparison with related plasmids from the *Bacillus cereus* group

Ariel F. Amadio [a,1], Graciela B. Benintende [a], Rubén O. Zandomeni [a,b,*]

[a] *Instituto de Microbiología y Zoología Agrícola (IMyZA), Instituto Nacional de Tecnología Agropecuaria (INTA), Las Cabañas y de Los Reseros, CP 1712 Castelar, Buenos Aires, Argentina*
[b] *Consejo Nacional de Investigaciones Científicas y Técnicas, Conicet, Argentina*

ARTICLE INFO

Article history:
Received 8 January 2009
Revised 28 July 2009
Available online 3 August 2009

Communicated by W. Klimke

Keywords:
Bacillus thuringiensis
Conjugation
Environmental isolate

ABSTRACT

Bacillus thuringiensis is an insect pathogen used worldwide as a bioinsecticide. It belongs to the *Bacillus cereus* sensu lato group as well as *Bacillus anthracis* and *B. cereus*. Plasmids from this group of organisms have been implicated in pathogenicity as they carry the genes responsible for different types of diseases that affect mammals and insects. Some plasmids, like pAW63 and pBT9727, encode a functional conjugation machinery allowing them to be transferred to a recipient cell. They also share extensive homology with the non-functional conjugation apparatus of pXO2 from *B. anthracis*. In this study we report the complete sequence of three plasmids from an environmental *B. thuringiensis* isolate from Argentina, obtained by a shotgun sequencing method. We obtained the complete nucleotide sequence of plasmids pFR12 (12 095 bp), pFR12.5 (12 459 bp) and pFR55 (55 712 bp) from *B. thuringiensis* INTA-FR7-4. pFR12 and pFR12.5 were classified as cryptic as they do not code for any obvious functions besides replication and mobilization. Both small plasmids were classified as RCR plasmids due to similarities with the replicases they encode. Plasmid pFR55 showed a structural organization similar to that observed for plasmids pAW63, pBT9727 and pXO2. pFR55 also shares a *tra* region with these plasmids, containing genes related to T4SS and conjugation. A comparison between pFR55 and conjugative plasmids led to the postulation that pFR55 is a conjugative plasmid. Genes related to replication functions in pFR55 are different to those described for plasmids with known complete sequences. pFR55 is the first completely sequenced plasmid with a replication machinery related to that of *ori*44. The analysis of the complete sequence of plasmids from an environmental isolate of *B. thuringiensis* permitted the identification of a near complete conjugation apparatus in pFR55, resembling those of plasmids pAW63, pBT9727 and pXO2. The availability of this sequence is a step forward in the study of the molecular basis of the conjugative process in Gram positive bacteria, particularly due to the similarity with known conjugation systems. It is also a contribution to the expansion of the non-pathogenic *B. cereus* plasmid gene pool.

© 2009 Elsevier Inc. All rights reserved.

1. Introduction

Members of the *Bacillus cereus* group of bacteria show pathological effects to different hosts. *Bacillus anthracis* is highly toxic to mammals and causes anthrax. *Bacillus thuringiensis* produces delta-endotoxin proteins that are toxic

* Corresponding author.
E-mail addresses: aamadio@cnia.inta.gov.ar (A.F. Amadio), gbenintende@cnia.inta.gov.ar (G.B. Benintende), zandomeni@cnia.inta.gov.ar (R.O. Zandomeni).
[1] Present address: Estación Experimental Agropecuaria Rafaela, Instituto Nacional de Tecnología Agropecuaria (INTA), Ruta 34 km 227 CC 22 CP2300, Rafaela, Santa Fe, Argentina.

0147-619X/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.plasmid.2009.07.005

to a large number of insect larvae and is the most commonly used biological pesticide in the world. *Bacillus cereus* is an opportunistic pathogen that is a common cause of food poisoning. Plasmids from the *B. cereus* group of organisms have been implicated in defining host range specificity and the pathogenicity (Rasko et al., 2005). The virulence of *B. anthracis* is due to the presence of two plasmids that encode the tripartite toxin and capsule (Okinaka et al., 1999a,b). *B. thuringiensis* is an insect pathogen, mainly due to the production of specific toxic proteins harbored mostly on plasmids (Schnepf et al., 1998). Although *B. cereus* was historically considered an opportunistic pathogen, several cases of gastrointestinal, periodontal and respiratory illnesses have been associated with *B. cereus* isolates (Helgason et al., 2000; Hoffmaster et al., 2004; Hoton et al., 2005; Rasko et al., 2007). Small plasmids from the *B. cereus* group are often referred to as cryptic, with few exceptions (Loeza-Lara et al., 2005; Fico and Mahillon, 2006). This is because the plasmids only code genes involved in their replication and mobilization (Andrup et al., 2003). Bigger plasmids are often implicated in pathogenicity and other marked phenotypic differences, but most functions of their genes remain unknown. A complete functional conjugative apparatus was described for *B. thuringiensis* pAW63 and pBT9727 plasmids (Van der Auwera et al., 2005, 2008), both closely related to the *B. anthracis* pXO2 plasmid. Since conjugation has been demonstrated for pAW63 in several media, including foodstuff (Van der Auwera et al., 2007), special attention should be given to these systems due to their potential impact on public health. As part of a nationwide screening program, several isolates of *B. thuringiensis* from Argentina were characterized (Benintende et al., 1999, 2000; Franco-Rivera et al., 2004). In this report we describe the complete sequence of three plasmids from the environmental isolate *B. thuringiensis* INTA-FR7-4 obtained by a shotgun sequencing strategy. We also analyze the common characteristics with other *B. cereus* group plasmids.

2. Materials and methods

2.1. Bacterial strains and culture conditions

Bacillus thuringiensis INTA-FR7-4 was isolated from a soil sample obtained in Misiones, Argentina, by using a selective and differential medium (PEMBA) Holbrook and Anderson, 1980. The strain was primarily identified by the presence of parasporal inclusion by using phase-contrast microscopy and characterized as *B. thuringiensis* subsp. *kumamotoensis* by using *B. cereus* group-specific repetitive extragenic palindromic sequence-based PCR analysis (data not shown). INTA-FR7-4 was grown on Terrific Broth (Sigma) at 30 °C overnight with vigorous agitation for alkaline lysis plasmid extraction. To perform CsCl gradient purification of plasmid DNA, INTA-FR7-4 was grown on Spizizen (Spizizen, 1958) medium at 30 °C overnight with vigorous agitation. *B. thuringiensis* subsp. *kurstaki* HD1 was used as a reference strain. *Escherichia coli* XL-1 blue strain (Stratagene), used for library transformation, was grown on Terrific Broth (Sigma) at 37 °C with the addition of ampicillin (50 μg/ml).

2.2. Plasmid DNA extraction and electrophoresis

Plasmid DNA from *B. thuringiensis* was obtained using two different techniques: an alkaline lysis procedure (modified from classical protocols) and a standard total DNA extraction followed by CsCl gradient purification (Sambrook et al., 1989). For the alkaline lysis method, a 100 ml overnight culture of *B. thuringiensis* was pelleted and washed twice with 25 ml TE (10 mM Tris–HCl, 1 mM EDTA, pH 8). Cells were re-suspended in 8 ml GTE (glucose 50 Mm, Tris–HCl 25 mM, EDTA 5 mM, pH8) + lysozyme (10 mg/ml), incubated at 37 °C for 30 min and stored on ice for 10 min. A fresh solution of 0.2 M NaOH/ 1% SDS was added (16 ml) and the suspension was incubated on ice until clarification occurred. Twelve milliliters of a 3 M sodium acetate solution was added and the suspension was centrifuged to recover the supernatant. The solution was incubated at 37 °C for 1 h following the addition of ribonuclease A up to 20 μg/ml. Next, two chloroform extractions were performed with gentle centrifugations to separate phases. One volume of isopropanol was added and the mixed solution was incubated at −70 °C for 30 min. Plasmid DNA was pelleted by centrifugation and washed with 70% ethanol. The dried pellet was dissolved in TE.

Electrophoresis of plasmid DNA was performed using a CHEF Mapper XA system (Bio-Rad). The parameters used for separation were determined using the autoalgorithm function with a 9–250 kb separation range. Electrophoresis was performed in a 1% PFC agarose (Bio-Rad) at 14 °C for 19 h.

For *E. coli* transformants, a classical plasmid extraction method was used (Sambrook et al., 1989).

2.3. Library construction

Approximately 1 μg of plasmid DNA from INTA-FR7-4 was diluted in one volume of glycerol 50% and nebulized at 8–10 psi N_2. Recovered DNA was purified by ethanol precipitation and electrophoresed in agarose 0.8%. Fragments ranging from 2 to 3 kb were purified using Sephaglass BandPrep Kit (Amersham) in accordance with the manufacturer's instructions. The purified fragments were repaired with Klenow fragment (Promega, USA) and ligated into pGEM3Z (Promega, USA) according to standard protocols (Sambrook et al., 1989).

2.4. Sequencing strategy

A shotgun sequencing strategy was used for the complete sequencing of plasmids (Fleischmann et al., 1995). Plasmid DNA from random clones was sequenced in an ABI377 (Applied Biosystems) using DYEnamic ET Terminator System (Amersham) and universal primers. Assembly was performed using a combination of *phred*/*phrap*/*Consed* (Ewing and Green, 1998; Ewing et al., 1998; Gordon et al., 1998) and GAP4 from Staden Package (Staden et al., 2000). Gap closure and finishing were achieved using a primer walking strategy for plasmids and PCR fragments of consecutive contigs. Primers were designed using a built-in

function of Gap4. Sequence verification was performed comparing *in silico* with real restriction analysis. The complete sequences of pFR12 (12 095 bp), pFR12.5 (12 459 bp) and pFR55 (55 712 bp) plasmids were determined. The sequences were deposited in GenBank under accession numbers [GenBank:EU362917], [GenBank:EU362918] and [GenBank:EU362919] respectively.

2.5. Sequence annotation and analysis

Potential coding regions were identified using a combination of GLIMMER v3 (Delcher et al., 2007) and EasyGene (Larsen and Krogh, 2003) both trained with the *B. cereus* ATCC 10987 genome (RefSeq:NC_003909) and a set of long ORFs from contigs that had been generated during the shotgun phase. Sequence similarities were determined using standalone BLAST programs (Altschul et al., 1997) to search nucleotide and non-redundant protein databases from GenBank. Comparisons against the InterPro database (Mulder et al., 2007) and Conserved Protein Domain database (Marchler-Bauer et al., 2007) were performed, using InterPro-scan (Quevillon et al., 2005) and BLAST respectively, in order to identify protein families and conserved domains. The annotation process was organized using Artemis v9 (Rutherford et al., 2000) and corrected manually gene-by-gene.

Circular diagrams of plasmids were created using GenomeDiagram (Pritchard et al., 2006). Comparisons among related plasmids were made using BLAST programs and visualized using Artemis Comparison Tool (ACT, Carver et al., 2005). Multiple sequence alignments were performed with ClustalX (Thompson et al., 1997).

3. Results

3.1. Plasmid profile from isolate INTA-FR7-4

Comparisons between the plasmid profiles of the *B. thuringiensis* isolate INTA-FR7-4 and the reference strain *B. thuringiensis* subsp. *kurstaki* HD1 (*Supplementary* Fig. 1) showed that the Argentine isolate has a more simple array of plasmids. This particular profile, and its lack of a relationship with type strains of the species, made the sequencing of extrachromosomal DNA present in INTA-FR7-4 particularly interesting.

3.2. General characteristics of plasmids

The complete nucleotide sequence of three plasmids from the INTA-FR7-4 isolate was determined using a shotgun sequencing method (Fleischmann et al., 1995). The shotgun results and main characteristics of the plasmids are listed in Table 1. There are two small plasmids of approximately 12 kb, and one 55 kb plasmid. The plasmids were named pFR12, pFR12.5 and pFR55 based on their sizes of 12 095, 12 459 and 55 712 bp respectively. A graphical representation of the plasmids is shown in Fig. 1. All of them have G + C contents of around 30%, a value that is similar to that of plasmids from related species, such as *B. anthracis* and *B. cereus* (Rasko et al., 2005; Andrup et al., 2003).

Restriction analysis with several endonucleases was performed as a control of the shotgun assembly process. All predicted restriction fragments were observed (data not shown).

3.3. Plasmid pFR12

The complete sequence of 12 095 bp contains 11 CDS's, eight of which (62.5%) are encoded on the same strand (*Supplementary* Table 1). Three CDS's encode hypothetical proteins and do not share significant similarities with database sequences. Similarities between the remaining CDS's are shown in *Supplementary* Table 1. RepFR12 has a 100% identity with the replicase from *B. mycoides* pDx14.2 (Di Franco et al., 2005) and a 68% identity with Rep14-3 from plasmid pTX14-3. RepFR12 belongs to group VII of RCR plasmid replicases (Andrup et al., 2003), as it is a 200 aminoacid Rep-protein with a double-strand origin (*dso*) located downstream from the *Rep*-gene. The replicase (RepFR12) similarities suggests that pFR12 replicates using a rolling circle replication (RCR) mechanism. *pFR12_01* codes a Mob protein (MobFR12) with a 92% identity with Mob from pDx14.2. pFR12 has the single-strand origin (*sso*) located upstream from the *mobFR12* gene, whereas the *dso* is located downstream from *repFR12* (Fig. 2A). Both the *dso* and *sso* were identified by alignment with corresponding regions of pTX14-3 and other related plasmids that belong to the same family. The organization of replication elements in pFR12, *dso-mob-rep-sso*, is very similar to that observed for pTX14-3, pDx14.2 and pE33L9 plasmids.

Three CDS's were found to have similarities with transcriptional regulators. pFR12_03 and pFR12_12 are similar to the transcriptional regulators from *B. cereus* E33L (pE33L9_0006) and several species of staphylococci. All these CDS's possess a *helix-turn-helix* (HTH) domain characteristic of the prokaryotic DNA-binding proteins from the xenobiotic response element (Xre) family of transcriptional regulators. pFR12_06 is similar to pSin9.7p03 (92% identity) and pXO2-68, both hypothetical proteins, and pE33L9_0008, which is a possible DNA-binding protein.

pFR12 shares a common backbone with pE33L9 from *B. cereus* E33L (Han et al., 2006) composed of replication and

Table 1
General characteristics of plasmids pFR12, pFR12.5 and pFR55. The G + C content are around 30%, characteristic of the species.

Plasmid	Size (bp)	% G + C	Genes number	Coding density	Average gene size (bp)*	% Coding	Coverage (×)
pFR12	12095	31.2	11	0.909	669	60.8	16.2
pFR12.5	12459	29.4	14	1.123	610	68.6	5.9
pFR55	55712	33.7	69	1.238	656	81.2	8.1

* The coding density is expressed in gene/kb.

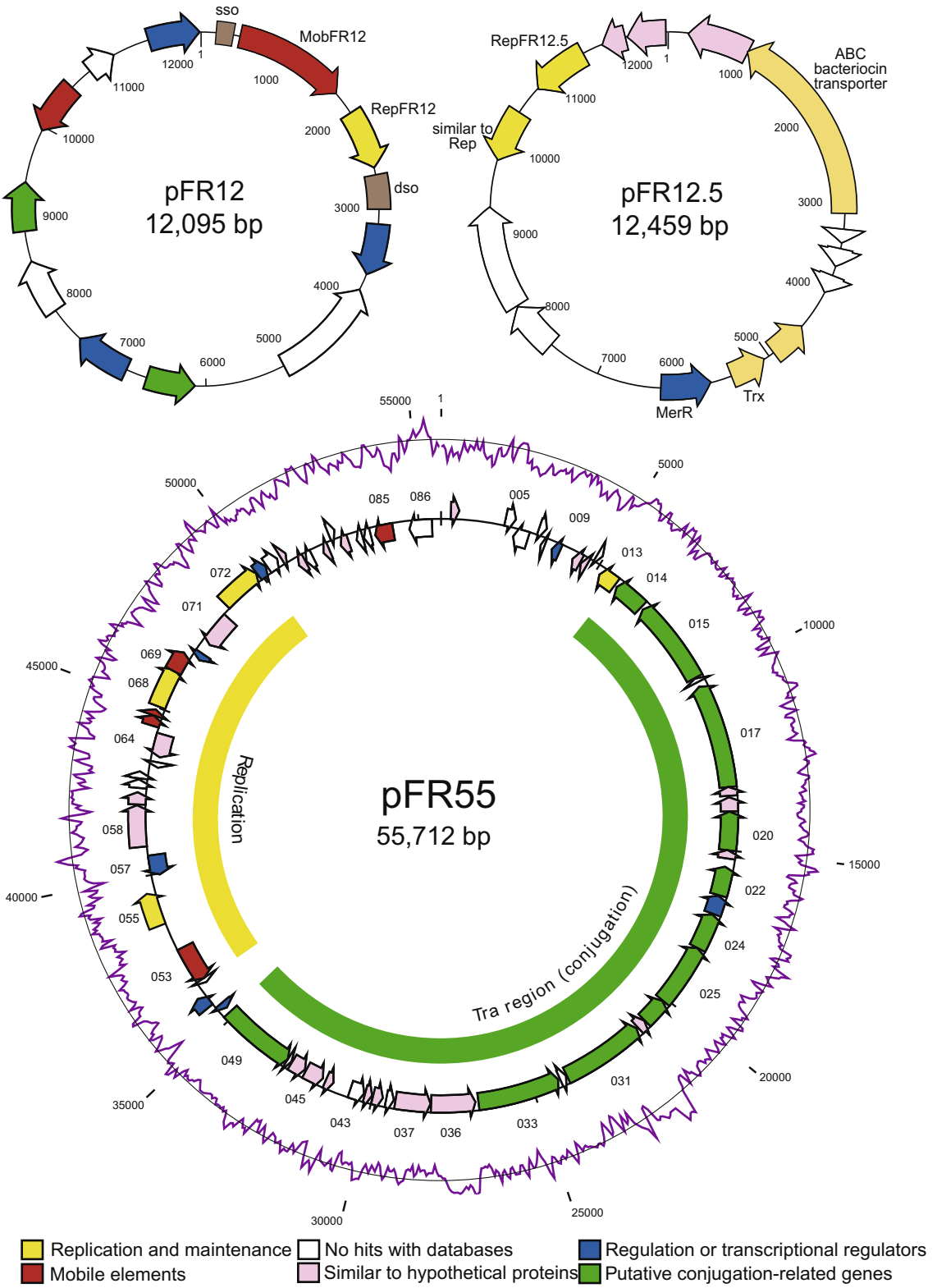

Fig. 1. Circular diagrams of plasmids pFR12, pFR12.5 and pFR55. The color of each CDS represents the functional categories described in the legend. In plasmid pFR55, regions responsible for replicative and conjugative processes are highlighted in the inner circle. The circles (from the center) represent genes encoded on each strand and G + C content. The CDS number corresponds to *Supplementary* Table 2.

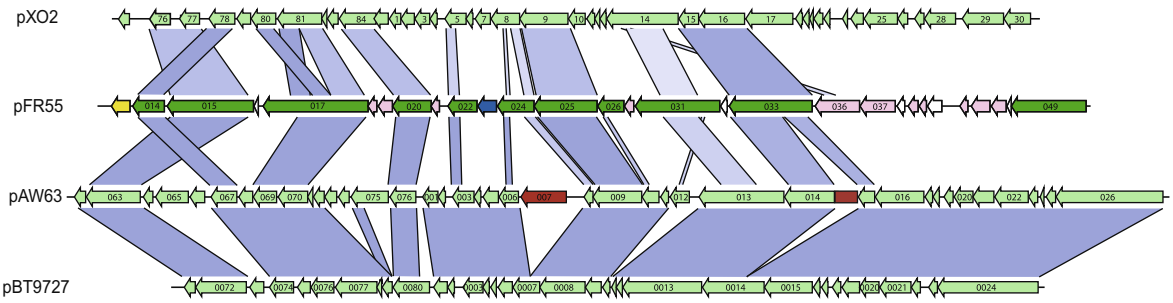

Fig. 2. Comparison of pFR55 *tra* region with related plasmids. Linear representation of the alignment of the *tra* region from plasmids pFR55, pXO2, pBT9727 and pWA63. Conserved regions are paired by shaded regions; the level of identity is indicated by the intensity of color and refers to the identity percentages in Table 2. The color of pFR55 features correspond to Fig. 2, as well as the CDS number indicated inside the arrows. Insertions interrupting CDS's in pAW63 are indicated in red and are absent in other plasmids. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

mobilization genes, a transcriptional regulator and an integrase/recombinase. It is also similar to the pSin9.7 plasmid from *B. mycoides* (Di Franco et al., 2005), although pSin9.7 has two copies of the transcriptional regulator and no integrase/recombinase (Supplementary Fig. 2).

3.4. Plasmid pFR12.5

The 12 459 bp sequence of plasmid pFR12.5 has 29.44% G + C content. Five of the fourteen CDS's share no similarities with database sequences (*Supplementary* Table 1), whereas three are similar to hypothetical proteins. As a result, protein function can only be inferred for 6 CDS's. Two proteins similar to replicases are encoded in pFR12.5. The first, pFR12.5_11 is similar to the predicted replicase from *B. cereus* pBC-Tim (89% identity) and Rep14-2 from *B. thuringiensis* pTX14-2. The second, pFR12.5_12, is similar to replicases from *B. cereus* pE33L9 (71% identity) and pE33L8 (55% identity), and Rep14-3 from *B. thuringiensis* pTX14-3. The two replicases from pFR12.5 exhibit only 34% identity and 53% similarity between them (*Supplementary* Fig. 3a), which are lower than the values obtained when comparing them with replicases from other plasmids (*Supplementary* Fig. 3b). Neither *dso* nor *sso* were identified in a conclusive way. However, the region downstream from pFR12.5_11 was similar to that of the *dso* from other small plasmids and could function as the double-stranded origin of replication in pFR12.5. These observations do not allow for a conclusive identification of the protein which acts as the replicase. A more detailed study is therefore required.

pFR12.5_02 is similar to proteins of the ABC-type bacteriocin transporter family from *B. thuringiensis* Al Hakam, *B. cereus* ATCC 10987, and the *Clostridium* species. It has an N-terminal peptidase domain, characteristic of bacteriocin transporters, followed by six transmembrane helices and the transporter itself.

The region consisting of genes *pFR12.5_13*, *pFR12.5_14*, *pFR12.5_01* and *pFR12.5_02* shows significant similarities with bacterial chromosomal genes from several related species. In the *B. cereus* ATCC 10987 genome, an ABC transporter and three hypothetical proteins are encoded in the same direction and order as those found in pFR12.5. This region is flanked by a transposase related to IS231 and an operon related to Tn-7. *B. thuringiensis* Al Hakam and *B. cereus* subsp. *cytotoxis* exhibits similar regions, particularly conserved in the transporter gene sequence.

3.5. Plasmid pFR55

The complete sequence of 55 715 bp from pFR55 has a 33.7% G + C content and 69 predicted genes, with an 81% coding density. Forty-eight (69.5%) of the predicted genes are hypothetical, 18 showed no hits against databases and the rest showed similarities with hypothetical proteins. As a result, function can only be assigned to 20 (30.5%) of the predicted proteins (Table 2 and *Supplementary* Table 2. Fifty of the 69 proteins (72.5%) in pFR55 are encoded on the same strand of the molecule, grouped as modules, a feature found in other plasmids of the genus (Van der Auwera et al., 2005). A characteristic modular organization is present in pFR55, with a *tra* region similar to those present in *B. thuringiensis* pAW63, *B. anthracis* pXO2 and *B. thuringiensis* pBT9727 plasmids. The genes related to replication and mobilization (Fig. 1) are collinear and grouped in different modules. pFR55 shows similarities with the replication region *ori*44 (Baum and Gilbert, 1991; Gamel and Piot, 1992). pFR55 does not share the common backbone (Van der Auwera et al., 2005) found in plasmids pXO2, pAW63 and pBT9727.

3.6. Replication

The protein proposed as responsible for the replication of pFR55 (RepFR55, pFR55_055) is similar to the functionally characterized replicase of the ~75 kDa plasmid from *B. thuringiensis* subsp. *kurstaki* HD73 (Gamel and Piot, 1992). However, there are no extended similarities between pFR55 replication regions and those from plasmids for which the complete sequence is available. Replicase from *ori*44 and RepFR55 share a 44% identity and 65% protein similarity. Both upstream and downstream regions are A + T rich, a characteristic shared with *ori*44 and other replicative regions (Huang et al., 2006). The repeats described downstream from the replicase of *ori*44 are not entirely conserved in pFR55 (*Supplementary* Fig. 4A). Two possible

Table 2
CDSs in *tra* region of pFR55: relevant similarities with databases and related plasmids.

CDS	Product	Position (start/ end)		Strand	Size (aa)	Most relevant blast hit (protein [species])	Identity (%)	Homolog in pXO1/pXO2	Homolog in pAW63	Homolog in pBT97-27
pFR55_014	Hypothetical protein	5810	6787	–	325	LtrC-like protein [*Bacillus thuringiensis* serovar *israelensis* ATCC 35646]	96	Hypothetical protein, pXO2-78	Hypothetical protein, pAW63_067	Hypothetical protein pBT9727_0074
pFR55_015	DNA topoisomerase, putative	6873	9530	–	885	DNA topoisomerase I [*Bacillus cereus* ATCC 10987]	70	topoisomerase I, pXO1-142		
pFR55_016	Hypothetical protein	9546	9680	–	44	–				
pFR55_017	Hypothetical protein	9826	13059	–	1077	RBTH_05426 [*Bacillus thuringiensis* serovar *israelensis* ATCC 35646]	90*	Hypothetical protein pXO2-80; hypothetical protein pXO2-81	Hypothetical protein pAW63_069; hypothetical protein pAW63_070	Hypothetical protein pBT9727_0076; hypothetical protein pBT9727_0077
pFR55_018	Hypothetical protein	13052	13327	–	91	Hypothetical protein RBTH_05430 [*Bacillus thuringiensis* serovar *israelensis* ATCC 35646]	95			
pFR55_019	Hypothetical protein	13370	13807	–	145	Hypothetical protein RBTH_05431 [*Bacillus thuringiensis* serovar *israelensis* ATCC 35646]	82			
pFR55_020	Hypothetical protein	13812	15002	–	396	Hypothetical cytosolic protein [*Bacillus thuringiensis* serovar *israelensis* ATCC 35646]	95	Hypothetical protein pXO2-84	Hypothetical protein pAW63_076	Hypothetical protein pBT9727_0080
pFR55_021	Hypothetical protein	14995	15255	–	86	Hypothetical protein BcerKBAB4DRAFT_3606 [*Bacillus weihenstephanensis* KBAB4]	88			
pFR55_022	Hypothetical protein	15515	16411	–	298	Hypothetical protein RBTH_05434 [*Bacillus thuringiensis* serovar *israelensis* ATCC 35646]	96	Hypothetical protein pXO2-05	Hypothetical protein pAW63_003	
pFR55_023	Transcriptional regulator, MerR family, putative	16433	17011	–	192	Transcriptional regulator, MerR family [*Bacillus thuringiensis* serovar *israelensis* ATCC 35646]	98			
pFR55_024	Lipoprotein, NLP/P60 family, putative	17039	18166	–	375	P54 protein precursor [*Bacillus thuringiensis* serovar *israelensis* ATCC 35646]	79*	Peptidase, M23/M37 family, pXO1-81; NLP/ P60 family protein, pXO2-08	Hypothetical protein pAW63_006-008	
pFR55_025	ATPase, Type IV secretion system, VirB4 component, putative	18171	20108	–	645	TRSE PROTEIN [*Bacillus thuringiensis* serovar *israelensis* ATCC 35646]	99	conserved hypothetical protein, pXO2-09	Hypothetical protein pAW63_009	Possible ATPase, pBT9727_0008
pFR55_026	Hypothetical protein	20124	20933	–	269	RBTH_02962 [*Bacillus thuringiensis* serovar *israelensis* ATCC 35646]	98*	Hypothetical protein pXO2-10	Hypothetical protein pAW63_010	Hypothetical protein pBT9727_0009
pFR55_028	Hypothetical protein	20948	21235	–	95	Hypothetical protein RBTH_02961 [*Bacillus thuringiensis* serovar *israelensis* ATCC 35646]	100			

(*continued on next page*)

Table 2 (*continued*)

CDS	Product	Position (start/end)	Strand	Size (aa)	Most relevant blast hit (protein [species])	Identity (%)	Homolog in pXO1/pXO2	Homolog in pAW63	Homolog in pBT97-27
pFR55_031	Membrane spanning protein, putative	21283 23883	–	866	Hypothetical membrane spanning protein [*Bacillus thuringiensis* serovar *israelensis* ATCC 35646]	95*	Membrane protein, pXO2-14	Hypothetical protein, pAW_013	
pFR55_032	Hypothetical protein	23907 24116	–	69	–				
pFR55_033	VirD4 homolog	24174 26729	–	851	ATPASE VIRD4 HOMOLOG [*Bacillus thuringiensis* serovar *israelensis* ATCC 35646]	91	Conjugation protein, TraG/TraD family, pXO2-16	Hypothetical protein pAW63_014-015	Possible conjugation protein, pBT9727_0014

* indicates that identity is calculated with only the aligned segment of proteins.

conformations for origin of replication are suggested: a shorter perfect repeat and a longer repeat with a mismatch (*Supplementary* Fig. 4). The latter would imply that perfect repeats are not needed, and that three-dimensional DNA conformation is sufficient for replicase recognition. Repeats upstream of *rep*FR55 are imperfect (*Supplementary* Fig. 4B), as are repeats in *ori*44, and are less conserved between them. Unlike other replication regions, no iterons are evident (Van der Auwera et al., 2005; Huang et al., 2006). Furthermore, there are no iterons in *ori*44, demonstrating that short repeats are not essential for replicase recognition.

pFR55 shows no similarities with genes involved in mechanisms such as copy number control or plasmid maintenance, as described by Huang et al. (2006). However, two genes adjacent to *repFR55* (*pFR55_053* and *pFR55_057*) have helix-turn-helix DNA-binding domains and might be involved in replication and maintenance functions. This assertion must be confirmed by further evidence. There are no homologs of plasmid maintenance systems in pFR55. Even so, *pFR55_013* encodes a restriction endonuclease, which may be involved in stability as restriction-methylation systems have been associated with a post-segregational host-killing model (Gerdes et al., 2000). However, there is no recognizable methylase associated with pFR55.

3.7. Analysis of tra region and comparison with related plasmids

The region between CDSs *pFR55_014* and *pFR55_033* is similar to those involved in the conjugation process described for plasmids pXO2, pAW63 and pBT9727 (Table 2 and Fig. 2). As is the case with those plasmids, the complete region is encoded in the same direction (defined as counterclockwise in the (Fig. 1). The region contains genes related to type IV secretion system (T4SS), which are responsible for the characteristic conjugative process of *Agrobacterium tumefaciens* (Yeo and Waksman, 2004; Christie, 2004). The presence of these types of genes does not guarantee functional conjugative transmission as pXO2 has a complete set of genes but is incapable of self-transmission (Reddy et al., 1987). In fact, pAW63, which is self-transmissible, has key components of the system interrupted by introns (Van der Auwera et al., 2005). Recently, pBT9727, which does not have any interrupted key genes, was proven to be self-transferable (Van der Auwera et al., 2008).

The list of components related to conjugative transfer encoded in pFR55 includes: a relaxase, a transglycosylase, a coupling factor and ATPases related to motor functions. *pFR55_015* encodes a topoisomerase similar to pAW63_063, topA of pBT9727 and a pseudogene of pXO2 (Fig. 2). These proteins are similar to the TrsI/TraI element from the plasmid pMRC01 from *Lactococcus lactis*, which has a relaxase activity in the *Lactococcus* conjugative system, which is equivalent to the role of VirD2 in *A. tumefaciens* conjugative machinery (Van der Auwera et al., 2005).

pFR55_017 encodes a conserved hypothetical protein present in the conjugative region of pXO2, pAW63 and pBT9727 with the particularity of having been "split" into two independent CDS's in these plasmids (Fig. 2). One of

those CDS's is a hypothetical protein, whereas the other belongs to the family of LtcR-like (*Lactococcus* transfer ORFC) putative conjugative elements. pFR55_014 is a protein similar to LtrC-like proteins, pAW63_67, pXO2-78 and pBT9727_0074 (Fig. 2). Although LtrC elements have been found in the *tra* regions of conjugative plasmids, their function has not been determined (Mills et al., 1994).

pFR55_024 has a domain similar to the NLP/P60 family of proteins with unknown function and a domain similar to transglycosylases and extracellular proteins related to invasion in *L. monocytogenes*. Its homolog in pAW63, pAW63_006, is interrupted by a group II intron named BthI1 (Van der Auwera et al., 2005) (Fig. 2). Interestingly, the interruption is present in pAW63, and absent in pBT9727, whereas both plasmids are self-transferable (Van der Auwera et al., 2005, 2008). pFR55_024 is also similar to pXO1-81. pFR55_025 is a VirB4 homolog of T4SS, conserved in pAW63 (pAW63_009), pBT9727 (pBT9727_0008) and pXO2 (pXO2-09). These proteins provide the energy for DNA transfer (Grohmann et al., 2003). pFR55_033 is a homolog of VirD4 ATPase present in pAW63, pBT9727 and pXO2. This gene is intact in pFR55 and pBT9727, but is interrupted in a different manner in pXO2 and pAW63 (Fig. 2) Van der Auwera et al., 2005. VirD4 is the coupling protein that connects the replication machinery, the transmembrane complex and the energy source (Grohmann et al., 2003). Recently TcpA, a homolog of FtsK/SpoIIIE present in plasmid pCW3 from *C. perfringes* was proposed as the coupling factor in the conjugative process (Parsons et al., 2007). Its role would be the same as that of VirD4, which is similar to TcpA and contains the same functional domains (Parsons et al., 2007). There is also a CDS similar to FtsK/SpoIIIE in pFR55 (pFR55_068) but it is outside the *tra* region.

3.8. Insertion sequences

Several genes share similarities with phage sequences (*Supplementary* Table 2). However, the presence of insertion sequences in pFR55 is less common than in other plasmids of the *B. cereus* group. A search for insertion sequences led to the identification of a segment similar to IS231 corresponding to region 55 545–55 649 bp. IS231 has been found in almost all members from the *B. cereus* group (De Palmenaer et al., 2004), and this partial hit suggests an ancestral event that inactivated the sequence.

3.9. Comparison with other plasmids

Besides the genes within the *tra* region, CDS's from pFR55 share similarities with different plasmids found in the *B. cereus* group. pFR55_009 is a transition state regulator AbrB homolog, also present in plasmid pXO1, pBT9727, pE33L54, pPER272 and in pBtoxis (two copies). AbrB may regulate a plasmid encoded process, since chromosomal and plasmid copies are different (Rasko et al., 2004) and could be an example of crosstalk between chromosome and plasmid DNA, although its function must be demonstrated.

pFR55_015 is a putative relaxase similar to pXO1-142 and topoisomerases from plasmids pBC10987, pE33L466 and pPER272. pFR55_024, a putative transglycosylase, is similar to pXO1-81. A recent bioinformatic study has indicated that pXO1 has a complete T4SS system which could be responsible for anthrax toxin secretion (Grynberg et al., 2007). pFR55_036 is similar to proteins encoded by pBC10987, pPER272 and pE44L466. These proteins are similar to adhesins with repeats resembling collagen. In *B. anthracis*, this protein is chromosomally encoded and is the immunodominant antigen of spore surface (Steichen et al., 2003; Sylvestre et al., 2002). Its function in pFR55 is unknown, but it could participate in cell-to-cell contact.

4. Discussion

The plasmid analysis of species belonging to the *B. cereus* group has become more important over the past few years due to its direct relationship with the pathogenic phenotype and the mobilization capabilities of these extrachromosomal elements (Hoffmaster et al., 2004; Rasko et al., 2005, 2007). It is crucial that the mechanisms implicated in replication, maintenance and gene transfer be understood, especially those involving toxin coding plasmids, which can affect public health.

The analysis of plasmid profiles from *B. thuringiensis* isolates is a technique used many times to describe isolates and correlate plasmid content with insecticidal activity (Lereclus et al., 1982; Ibarra and Federici, 1987; Aptosoglou et al., 1997). The plasmid profile of type strain has recently been used to identify them, using an extraction method similar to the one used here, but with regular instead of pulsed field gel electrophoresis (Reyes-Ramirez and Ibarra, 2008). Unfortunately, the use of regular electrophoresis does not permit the analysis of plasmids migrating above the chromosomal DNA band, limiting the analysis to just small plasmids. Additional recent studies have described a method to obtain larger plasmids from the *B. cereus* group (Andrup et al., 2008). This method allows for the detection of plasmids, but does not produce quality DNA for further analysis. In this present study we describe an alternative protocol for purifying plasmid DNA suitable for library construction or restriction analysis which involves the use of pulsed field gel electrophoresis to separate and distinguish plasmids of all sizes.

Several genome projects have yielded the complete genome sequences from different species of mostly pathogenic *B. cereus* group organisms, although non-pathogenic isolates are still underrepresented in the gene pool (Rasko et al., 2005). Here, we have determined the complete sequence of three plasmids from an environmental isolate of *B. thuringiensis* using a shotgun sequencing strategy. Two plasmid libraries were constructed, one with a CsCl gradient purification of total DNA and a second one based on an alkaline lysis. The shotgun process began with the first library but was overrepresented with pFR12 plasmid, as can be observed in the coverage (Table 1). This led to the development of the second library using plasmid DNA obtained by the alkaline lysis method. The use of both libraries showed a more adequate representation of contigs and permitted the closure of three plasmid sequences.

Most of the proteins predicted for the isolate INTA-FR7-4 are hypothetical (62%), whether they are similar to hypo-

thetical proteins or they do not show hits against databases. pFR55 shows the highest number of hypothetical proteins, with 48 of 69 CDS's.

Small plasmids from *B. cereus* group organisms encode only a few proteins, and the majority of their functions are related to plasmid replication and mobilization (Andrup et al., 2003). Plasmid pFR12 may replicate via RCR, as RepFR12 is very similar to other replicases from related plasmids. The region involved in replication and mobilization is shared by a group of plasmids, including pE33L9 and pSin9.7, with a common backbone (*Supplementary* Fig. 2), and pTX14-3, pDx14.2 and pE33L8. All plasmids share (with small differences) the *dso* and *sso*, indicating that the block has led to the formation of a family of plasmids with different accompanying genes; these genes make each plasmid unique. The mobility of such blocks could be the reason for the presence of two *rep*-genes in pFR12.5. The degree of similarity between the two pFR12.5 Rep proteins is lower than the similarity observed with other replicases (*Supplementary* Fig. 3), indicating a distinct origin or an ancestral insertion event.

pFR12.5 encodes two proteins involved in protein modification by disulfide bond formation. Trx and DsaB proteins are generally chromosomally encoded, and the function of these genes in the plasmid must be analyzed. In addition to this, the presence of a region conserved in chromosomal genomes suggests that gene transfer between chromosomal and plasmid DNA has occurred. This is supported by the fact that this region is absent in some of the *B. cereus* group genomes.

The modular structure of pFR55 resembles other plasmids from the *B. cereus* group, like pAW63, pBT9727 and pXO2 (Van der Auwera et al., 2005). While they do not share a common backbone, the *tra* region is conserved. The replication system of pFR55 is different from any of the completely sequenced plasmids. The only similar machinery has been described for *ori*44 (Baum and Gilbert, 1991), but the complete sequence of the plasmid is not available in public databases. Replicases from pFR55 and *ori*44 are similar to each other but significantly different to other plasmid replicases, suggesting a different mechanism of replication. Inverted repeats up- and downstream from RepFR55 show differences to those of *ori*44 (*Supplementary* Fig. 4). A perfect inverted repeat is found at ~1200 bp from RepFR55, near the next CDS (pFR55_056). pFR55_056 shows a DNA-binding domain and could be involved in replication or maintenance. There is no significant similarity between other replicative systems and genes in pFR55, so further characterization of pFR55 replication mechanisms should be made. Currently, all characterized replicons from *B. cereus* group plasmids belong to the pAMβ1 family of theta replicons. A more detailed analysis must be performed to determine whether pFR55 replicates in a similar way, or whether it corresponds to another family of theta plasmids.

The *tra* region of pFR55 is similar to that of plasmids pAW63, pBT9727 and pXO2, and also resembles regions from *Enterococcus*, *Streptococcus*, and *Clostridium* plasmids involved in the spreading of antibiotic resistance determinants (Grohmann et al., 2003). The current model of conjugative transfer proposes that two protein complexes, the relaxosome and the mating pair formation (mpf) complex, are involved in the process. The mpf complex is ancestrally related to T4SS (Christie, 2001). T4SS comprises a family of transporters involved in several processes, including conjugative DNA transfer and virulence. *A. tumefaciens* plasmid transfer to the plant cell could be considered a combination of conjugation and virulence, since conjugative DNA transfer produces plant disease (de Paz et al., 2005). Among the comparable components present in pFR55, pFR55_015 is similar to VirD2/TraI, which acts as a relaxase. The relaxase catalyzes the cleavage of plasmid at the *nic* site of *ori*T, initiating the conjugation process (Grohmann et al., 2003).

Plasmid pFR55 encodes proteins similar to T4SS components, as well as other *B. cereus* group plasmids. Among the components present in pFR55 are: pFR55_033, a VirD4/TraG homolog, possibly the coupling factor connecting the relaxosome to the translocation apparatus. pFR55_068, an FtsK/SpoIII/TcpA homolog, may also function as the coupling factor if the conjugation machinery in pFR55 is active. Further studies must be carried out to determine whether pFR55_033 or pFR55_068 might function as coupling proteins acting by themselves or in combination.

Another component of T4SS present in pFR55 is pFR55_024, a VirB1 homolog. This acts as a transglycosylase which produces the local opening of peptidoglycan layer, allowing the DNA–protein complex to pass across the cell envelope (Grohmann et al., 2003). pFR55_025, a homolog of VirB4/TraL ATPase, which could function as a signal transducer possibly inducing conformational changes (Grohmann et al., 2003), is located downstream of pFR55_024.

Similarities observed between plasmids from the *B. cereus* group and Gram negative conjugative plasmids correspond to the internal components of the conjugative system associated with the internal membrane. Specific external components should manage the main variations in structure and functions such as the thin peptidoglycan layer and the absence of an external membrane (Van der Auwera et al., 2005). Cell-to-cell interaction could be resolved in pFR55 with the presence of several membrane proteins. pFR55_031 is a membrane protein, and pFR55_036 and pFR55_049 are similar to membrane and collagen adhesion proteins and could be involved in such interactions.

An absent component in pFR55, when compared to other conjugation-related plasmids, is a homolog of VirB11, which is the putative chaperone involved in translocation through the membrane system (Grohmann et al., 2003).

A global analysis of pFR55, and its comparison with related plasmids, suggests that it is a conjugative plasmid, although its conjugation capability remains to be shown. The complete sequence of this plasmid will permit further experiments which will enable us to gain a better understanding of the role of each component in the conjugative process of Gram positive plasmids, on a molecular level.

5. Conclusions

In this present study we have described the complete sequence of three plasmids from an environmental *B. thuringiensis* isolate.

The two small plasmids (of approximately 12 kb) have a limited set of genes, mostly related to plasmid replication and transfer. Plasmid pFR55 possesses a modular structure including a conjugation-related region similar to the conjugative plasmids previously described. This *tra* region resembles those from plasmids pAW63, pBT9727 and pXO2 which all have their genes encoded on the same strand. With the exception of the *tra* region, no extended similarities were found with other completely sequenced plasmids. Replication machinery of pFR55 is similar to that previously described for *ori*44. The availability of the complete sequence of plasmid pFR55 permits genetic studies related to establishing the individual role of components in the conjugation process and also helps to increase the number of completely sequenced and annotated plasmids, especially those from the non-pathogenic *B. cereus* group species.

6. GenBank

The sequences of pFR12, pFR12.5 and pFR55 were deposited in GenBank under accession numbers [GenBank: EU362917], [GenBank:EU362918] and [GenBank:EU362919] respectively.

Acknowledgments

We gratefully thank Nanci Lopez, Carmen Mercado and Irma Fuxan for their technical support in sequencing and routine laboratory protocols. A.A. performed this work with a fellowship from CONICET. The work was funded by project INTA-AETA1672.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.plasmid.2009.07.005.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Andrup, L., Jensen, G.B., Wilcks, A., Smidt, L., Hoflack, L., Mahillon, J., 2003. The patchwork nature of rolling-circle plasmids: comparison of six plasmids from two distinct *Bacillus thuringiensis* serotypes. Plasmid 49, 205–232.

Andrup, L., Barfod, K.K., Jensen, G.B., Smidt, L., 2008. Detection of large plasmids from the *Bacillus cereus* group. Plasmid 59, 139–143.

Aptosoglou, S.G., Sivropoulou, A., Koliais, S.I., 1997. Plasmid patterns of *Bacillus thuringiensis* strains and isolates. Microbios 91, 203–214.

Baum, J.A., Gilbert, M.P., 1991. Characterization and comparative sequence analysis of replication origins from three large *Bacillus thuringiensis* plasmids. J. Bacteriol. 173, 5280–5289.

Benintende, G.B., López-Meza, J.E., Cozzi, J.G., Ibarra, J.E., 1999. Novel non-toxic isolates of *Bacillus thuringiensis*. Lett. Appl. Microbiol. 29, 151–155.

Benintende, G.B., Lopez-Meza, J.E., Cozzi, J.G., Piccinetti, C.F., Ibarra, J.E., 2000. Characterization of INTA 51-3, a new atypical strain of *Bacillus thuringiensis* from Argentina. Curr. Microbiol. 41, 396–401.

Carver, T.J., Rutherford, K.M., Berriman, M., Rajandream, M.A., Barrell, B.G., Parkhill, J., 2005. ACT: the Artemis Comparison Tool. Bioinformatics 21, 3422–3423.

Christie, P.J., 2001. Type IV secretion: intercellular transfer of macromolecules by systems ancestrally related to conjugation machines. Mol. Microbiol. 40, 294–305.

Christie, P.J., 2004. Type IV secretion: the Agrobacterium VirB/D4 and related conjugation systems. Biochim. Biophys. Acta 1694, 219–234.

De Palmenaer, D., Vermeiren, C., Mahillon, J., 2004. IS231-MIC231 elements from *Bacillus cereus* sensu lato are modular. Mol. Microbiol. 53, 457–467.

de Paz, H.D., Sangari, F.J., Bolland, S., Garcia-Lobo, J.M., Dehio, C., de la Cruz, F., Llosa, M., 2005. Functional interactions between type IV secretion systems involved in DNA transfer and virulence. Microbiology 151, 3505–3516.

Delcher, A.L., Bratke, K.A., Powers, E.C., Salzberg, S.L., 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics 23, 673–679.

Di Franco, C., Santini, T., Pisaneschi, G., Beccari, E., 2005. Insights into the genetic organization of the *Bacillus mycoides* cryptic plasmids pDx14.2 and pSin9.7 deduced from their complete nucleotide sequence. Plasmid 54, 288–293.

Ewing, B., Green, P., 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. 8, 186–194.

Ewing, B., Hillier, L., Wendl, M.C., Green, P., 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. 8, 175–185.

Fico, S., Mahillon, J., 2006. TasA–tasB, a new putative toxin–antitoxin (TA) system from *Bacillus thuringiensis* pGI1 plasmid is a widely distributed composite mazE-doc TA system. BMC Genomics 7, 259.

Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.F., Dougherty, B.A., Merrick, J.M., et al, 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science 269, 496–512.

Franco-Rivera, A., Benintende, G., Cozzi, J., Baizabal-Aguirre, V.M., Valdez-Alarcon, J.J., Lopez-Meza, J.E., 2004. Molecular characterization of *Bacillus thuringiensis* strains from Argentina. Antonie Van Leeuwenhoek 86, 87–92.

Gamel, P.H., Piot, J.C., 1992. Characterization and properties of a novel plasmid vector for *Bacillus thuringiensis* displaying compatibility with host plasmids. Gene 120, 17–26.

Gerdes, K., Ayora, S., Canosa, I., Ceglowski, P., Diaz-Orejas, R., Franch, T., Gultyaev, A.P., Bugge Jensen, R., Kobayashi, I., Macpherson, C., Summers, D., Thomas, C.M., Zielenkiewicz, U., 2000. Plasmid maintenance systems. In: Thomas, C.M. (Ed.), The Horizontal Gene Pool. Bacterial Plasmids and Gene Spread. Harwood Academic Publishers, Amsterdam, pp. 49–86.

Gordon, D., Abajian, C., Green, P., 1998. Consed: a graphical tool for sequence finishing. Genome Res. 8, 195–202.

Grohmann, E., Muth, G., Espinosa, M., 2003. Conjugative plasmid transfer in gram-positive bacteria. Microbiol. Mol. Biol. Rev. 67, 277–301 (table of contents).

Grynberg, M., Li, Z., Szczurek, E., Godzik, A., 2007. Putative type IV secretion genes in *Bacillus anthracis*. Trends Microbiol. 15, 191–195.

Han, C.S., Xie, G., Challacombe, J.F., Altherr, M.R., Bhotika, S.S., Brown, N., Bruce, D., Campbell, C.S., Campbell, M.L., Chen, J., Chertkov, O., Cleland, C., Dimitrijevic, M., Doggett, N.A., Fawcett, J.J., Glavina, T., Goodwin, L.A., Green, L.D., Hill, K.K., Hitchcock, P., Jackson, P.J., Keim, P., Kewalramani, A.R., Longmire, J., Lucas, S., Malfatti, S., McMurry, K., Meincke, L.J., Misra, M., Moseman, B.L., Mundt, M., Munk, A.C., Okinaka, R.T., Parson-Quintana, B., Reilly, L.P., Richardson, P., Robinson, D.L., Rubin, E., Saunders, E., Tapia, R., Tesmer, J.G., Thayer, N., Thompson, L.S., Tice, H., Ticknor, L.O., Wills, P.L., Brettin, T.S., Gilna, P., 2006. Pathogenomic sequence analysis of *Bacillus cereus* and *Bacillus thuringiensis* isolates closely related to *Bacillus anthracis*. J. Bacteriol. 188, 3382–3390.

Helgason, E., Caugant, D.A., Olsen, I., Kolsto, A.B., 2000. Genetic structure of population of *Bacillus cereus* and *B. thuringiensis* isolates associated with periodontitis and other human infections. J. Clin. Microbiol. 38, 1615–1622.

Hoffmaster, A.R., Ravel, J., Rasko, D.A., Chapman, G.D., Chute, M.D., Marston, C.K., De, B.K., Sacchi, C.T., Fitzgerald, C., Mayer, L.W., Maiden, M.C., Priest, F.G., Barker, M., Jiang, L., Cer, R.Z., Rilstone, J., Peterson, S.N., Weyant, R.S., Galloway, D.R., Read, T.D., Popovic, T., Fraser, C.M., 2004. Identification of anthrax toxin genes in a *Bacillus cereus* associated with an illness resembling inhalation anthrax. Proc. Natl. Acad. Sci. USA 101, 8449–8454.

Holbrook, R., Anderson, J.M., 1980. An improved selective and diagnostic medium for the isolation and enumeration of *Bacillus cereus* in foods. Can. J. Microbiol. 26, 753–759.

Hoton, F.M., Andrup, L., Swiecicka, I., Mahillon, J., 2005. The cereulide genetic determinants of emetic *Bacillus cereus* are plasmid-borne. Microbiology 151, 2121–2124.

Huang, J., Guo, S., Mahillon, J., Van der Auwera, G.A., Wang, L., Han, D., Yu, Z., Sun, M., 2006. Molecular characterization of a DNA fragment harboring the replicon of pBMB165 from *Bacillus thuringiensis* subsp. *tenebrionis*. BMC Genomics 7, 270.

Ibarra, J.E., Federici, B.A., 1987. Comparison of the toxicity, parasporal body protein composition, and plasmid complements of nine isolates of *Bacillus thuringiensis* subsp. *israelensis*. J. Econ. Entomol. 80, 1132–1136.

Larsen, T.S., Krogh, A., 2003. EasyGene – a prokaryotic gene finder that ranks ORFs by statistical significance. BMC Bioinf. 4, 21.

Lereclus, D., Lecadet, M.M., Ribier, J., Dedonder, R., 1982. Molecular relationships among plasmids of *Bacillus thuringiensis*: conserved sequences through 11 crystalliferous strains. Mol. Gen. Genet. 186, 391–398.

Loeza-Lara, P.D., Benintende, G., Cozzi, J., Ochoa-Zarzosa, A., Baizabal-Aguirre, V.M., Valdez-Alarcon, J.J., Lopez-Meza, J.E., 2005. 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.

Marchler-Bauer, A., Anderson, J.B., Derbyshire, M.K., DeWeese-Scott, C., Gonzales, N.R., Gwadz, M., Hao, L., He, S., Hurwitz, D.I., Jackson, J.D., Ke, Z., Krylov, D., Lanczycki, C.J., Liebert, C.A., Liu, C., Lu, F., Lu, S., Marchler, G.H., Mullokandov, M., Song, J.S., Thanki, N., Yamashita, R.A., Yin, J.J., Zhang, D., Bryant, S.H., 2007. CDD: a conserved domain database for interactive domain family analysis. Nucleic Acids Res. 35, D237–D240.

Mills, D.A., Choi, C.K., Dunny, G.M., McKay, L.L., 1994. Genetic analysis of regions of the *Lactococcus lactis* subsp. *lactis* plasmid pRS01 involved in conjugative transfer. Appl. Environ. Microbiol. 60, 4413–4420.

Mulder, N.J., Apweiler, R., Attwood, T.K., Bairoch, A., Bateman, A., Binns, D., Bork, P., Buillard, V., Cerutti, L., Copley, R., Courcelle, E., Das, U., Daugherty, L., Dibley, M., Finn, R., Fleischmann, W., Gough, J., Haft, D., Hulo, N., Hunter, S., Kahn, D., Kanapin, A., Kejariwal, A., Labarga, A., Langendijk-Genevaux, P.S., Lonsdale, D., Lopez, R., Letunic, I., Madera, M., Maslen, J., McAnulla, C., McDowall, J., Mistry, J., Mitchell, A., Nikolskaya, A.N., Orchard, S., Orengo, C., Petryszak, R., Selengut, J.D., Sigrist, C.J., Thomas, P.D., Valentin, F., Wilson, D., Wu, C.H., Yeats, C., 2007. New developments in the InterPro database. Nucleic Acids Res. 35, D224–D228.

Okinaka, R., Cloud, K., Hampton, O., Hoffmaster, A., Hill, K., Keim, P., Koehler, T., Lamke, G., Kumano, S., Manter, D., Martinez, Y., Ricke, D., Svensson, R., Jackson, P., 1999a. Sequence, assembly and analysis of pX01 and pX02. J. Appl. Microbiol. 87, 261–262.

Okinaka, R.T., Cloud, K., Hampton, O., Hoffmaster, A.R., Hill, K.K., Keim, P., Koehler, T.M., Lamke, G., Kumano, S., Mahillon, J., Manter, D., Martinez, Y., Ricke, D., Svensson, R., Jackson, P.J., 1999b. Sequence and organization of pXO1, the large *Bacillus anthracis* plasmid harboring the anthrax toxin genes. J. Bacteriol. 181, 6509–6515.

Parsons, J.A., Bannam, T.L., Devenish, R.J., Rood, J.I., 2007. TcpA, an FtsK/SpoIIIE homolog, is essential for transfer of the conjugative plasmid pCW3 in *Clostridium perfringens*. J. Bacteriol. 189, 7782–7790.

Pritchard, L., White, J.A., Birch, P.R., Toth, I.K., 2006. GenomeDiagram: a python package for the visualization of large-scale genomic data. Bioinformatics 22, 616–617.

Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., Lopez, R., 2005. InterProScan: protein domains identifier. Nucleic Acids Res. 33, W116–W120.

Rasko, D.A., Ravel, J., Okstad, O.A., Helgason, E., Cer, R.Z., Jiang, L., Shores, K.A., Fouts, D.E., Tourasse, N.J., Angiuoli, S.V., Kolonay, J., Nelson, W.C., Kolsto, A.B., Fraser, C.M., Read, T.D., 2004. The genome sequence of *Bacillus cereus* ATCC 10987 reveals metabolic adaptations and a large plasmid related to *Bacillus anthracis* pXO1. Nucleic Acids Res. 32, 977–988.

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

Rasko, D.A., Rosovitz, M.J., Okstad, O.A., Fouts, D.E., Jiang, L., Cer, R.Z., Kolsto, A.B., Gill, S.R., Ravel, J., 2007. Complete sequence analysis of novel plasmids from emetic and periodontal *Bacillus cereus* isolates reveals a common evolutionary history among the *B. cereus*-group plasmids, including *Bacillus anthracis* pXO1. J. Bacteriol. 189, 52–64.

Reddy, A., Battisti, L., Thorne, C.B., 1987. Identification of self-transmissible plasmids in four *Bacillus thuringiensis* subspecies. J. Bacteriol. 169, 5263–5270.

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

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

Sambrook, J., Fitsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Press, Cold Spring Harbor.

SantaLucia Jr., J., 1998. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. Proc Natl Acad Sci USA 95, 1460–1465.

Schnepf, E., Crickmore, N., Van Rie, J., Lereclus, D., Baum, J., Feitelson, J., Zeigler, D.R., Dean, D.H., 1998. *Bacillus thuringiensis* and its pesticidal crystal proteins. Microbiol. Mol. Biol. Rev. 62, 775–806.

Spizizen, J., 1958. Transformation of biochemically deficient strains of *Bacillus subtilis* by deoxyribonucleate. Proc. Natl. Acad. Sci. USA 44, 1072–1078.

Staden, R., Beal, K.F., Bonfield, J.K., 2000. The Staden package, 1998. Methods Mol Biol. 132, 115–130.

Steichen, C., Chen, P., Kearney, J.F., Turnbough Jr., C.L., 2003. Identification of the immunodominant protein and other proteins of the *Bacillus anthracis* exosporium. J. Bacteriol. 185, 1903–1910.

Sylvestre, P., Couture-Tosi, E., Mock, M., 2002. A collagen-like surface glycoprotein is a structural component of the *Bacillus anthracis* exosporium. Mol. Microbiol. 45, 169–178.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.

Van der Auwera, G.A., Andrup, L., Mahillon, J., 2005. 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 6, 103.

Van der Auwera, G.A., Timmery, S., Hoton, F., Mahillon, J., 2007. Plasmid exchanges among members of the *Bacillus cereus* group in foodstuffs. Int. J. Food Microbiol. 113, 164–172.

Van der Auwera, G.A., Timmery, S., Mahillon, J., 2008. Self-transfer and mobilisation capabilities of the pXO2-like plasmid pBT9727 from *Bacillus thuringiensis* subsp. *konkukian* 97-27. Plasmid 59, 134–138.

Yeo, H.J., Waksman, G., 2004. Unveiling molecular scaffolds of the type IV secretion system. J. Bacteriol. 186, 1919–1926.