

## Draft genome sequence of *Bacillus thuringiensis* strain S1287, an isolate showing insecticidal activity against coleoptera

### Projeto de sequência do genoma de *Bacillus thuringiensis* strain S1287, um isolado mostrando actividade insecticida contra coleoptera

DOI:10.34115/basrv6n3-009

Recebimento dos originais: 18/02/2022

Aceitação para publicação: 14/03/2022

#### **Paulo Queiroz**

Doutorado em Biologia Animal pela Universidade de Brasília (UNB)

Institution: Centro Universitário de Brasília (CEUB)

Address: 707/907, Campus Universitário, Asa Norte, Brasília, DF, CEP: 70790-075

E-mail: prqzqueiroz@gmail.com

#### **Roberto Togawa**

Doutorado pela University of Bedfordshire

Institution: Embrapa Recursos Genéticos e Biotecnologia

Address: Parque Estação Biológica, PqEB, Av. W5 Norte (final) Caixa Postal 02372, Brasília, DF, CEP: 70770-917

E-mail: roberto.togawa@embrapa.br

#### **Érica Martins**

Doutorado em Ciências Biológicas (Biologia Molecular) pela Universidade de Brasília

Institution: Universidade de Brasília

Address: SBN Quadra 02 Bloco C – Edifício Phenícia, CEP: 70040-020

E-mail: es.martins2@gmail.com

#### **Priscila Grynberg**

Doutorado em Bioinformática pela Universidade Federal de Minas Gerais (UFMG)

Institution: Embrapa Recursos Genéticos e Biotecnologia

Address: Parque Estação Biológica, PqEB, Av. W5 Norte (final) Caixa Postal 02372, Brasília, DF, CEP: 70770-917

E-mail: priscila.grynberg@embrapa.br

#### **Rose Monnerat**

Doutorado em Agronomia na Ecole Nationale Agronomique de Montpellier

Institution: Embrapa Recursos Genéticos e Biotecnologia

Address: Parque Estação Biológica, PqEB, Av. W5 Norte (final) Caixa Postal 02372, Brasília, DF, CEP: 70770-917

E-mail: rose.monnerat@embrapa.br

#### **ABSTRACT**

*Bacillus thuringiensis* is an important bacterium showing insecticide proteins effective to control several important medical pests, including several Diptera species, like Diptera, Coleoptera and Lepidoptera. This work describes the draft genome sequence of *B. thuringiensis* S1287, which contains genes encoding the parasporal crystal cry, like Cry1Ab, Cry1Bb, Cry1C, Cry1D, Cry1Fb, Cry1Hb, Cry1Id, Cry1Ja, Cry1Ka and

Cry1Nn. The genes belonging to the Vip Family were identified, Vpb1Aa, Vpa2Aa and Vip3Ag. Genes coding for plant growth promotion pathway were found.

**Keywords:** insecticidal proteins, cry genes, boll weevil.

## RESUMO

*Bacillus thuringiensis* é uma bactéria importante mostrando proteínas inseticidas eficazes para controlar várias pragas agrícolas importantes, incluindo várias espécies de Diptera, como a Diptera, Coleoptera e Lepidoptera. Este trabalho descreve o projecto da sequência genómica de *B. thuringiensis* S1287, que contém genes que codificam o grilo de cristal parasporal, como Cry1Ab, Cry1Bb, Cry1C, Cry1D, Cry1Fb, Cry1Hb, Cry1Id, Cry1Ja, Cry1Ka e Cry1Nn. Foram identificados os genes pertencentes à Família Vip, Vpb1Aa, Vpa2Aa e Vip3Ag. Foram encontrados genes que codificam a via de promoção do crescimento das plantas.

**Palavras-chave:** proteínas inseticidas, genes de choro, gorgulho de boll.

## 1 INTRODUCTION

*Bacillus thuringiensis* is a soil bacterium of considerable economic importance used as a biocontrol agent (1), including the cotton Boll weevil (*Anthonomus grandis*).

The toxic action of some *B. thuringiensis* strains is due the parasporal inclusions that are composed of several toxic polypeptides named Cry toxins. And Vip toxins which can act together to affect the coleopteran larvicidal activity (2).

Purified genomic and plasmid DNA's from strain S1287 were extracted by Masterpure Gram-positive DNA purification kit (Epicentre) and QIAGEN Plasmid Maxi Kit (QiAGEN). Afterwards, these DNA's were sequenced at Macrogen, Inc. (Seoul, Korea) using high-throughput HiSeq2000 and GS-FLX Plus platforms getting one lane of 100 bp and 1/8 region plate, respectively. The reads were assembled using SOAP de novo (version 1.05) and produced 218 contigs totaling 6,636,210 bp (Q20 = 99.34%), with a maximum scaffold size of 359,449 bp, an N<sub>50</sub> length of 115,122 bp, and 34.22% G+C content and genome coverage depth was approximately 100x. De novo assembly was carried out using the Geneious version (8.0.4) (3). A total of 6,554 coding sequences (4), 71 tRNA (5), 5 rRNA operons (6) and 5 ncRNA were predicted. We identified 5 plasmids ranging from 342,958 to 8,235 bp size. The functions of encoding genes were annotated by using the NCBI nr, Swiss-Prot (7), Clusters of Orthologous Groups (COG) (8), KEGG (9), and InterProScan (10) databases.

The custom insecticidal toxin database was constructed with nucleotide and amino acid sequences of *cry*, *vip*, *cyt*, growth promotion, parasporin and bacteriocin genes using

the complete sequences of other *B. thuringiensis* strains deposited in public databases (<http://www.ncbi.nlm.nih.gov/genome/> or <http://www.bpprc.org/>). The strain S1785 draft genome sequence carries several insecticidal toxin genes showing identities to the *CryIAb*, *CryIBb*, *CryIC*, *CryID*, *CryIFb*, *CryIHb*, *CryIId*, *CryIJa*, *CryIKa* and *CryINn*. and the *Vip* family genes *Vpb1Aa*, *Vpa2Aa* and *Vip3Ag*. The genes which encode the metabolic pathway of plant growth promotion 1-aminocyclopropane-1-carboxylate deaminase, indolepyruvate decarboxylase, putative acid phosphatase and siderophore biosynthesis protein were identified. Using the new nomenclature of proteins toxic to insects proposed by Crickmore et al. (11) we found the *Spp1Aa* gene. The genomic analyses could not identify any operon related to the antimicrobial peptide bacteriocin (thuricin or nisin).

**Nucleotide sequence accession numbers.** The sequence of the *B. thuringiensis* strain S1287 has been deposited in GenBank with the accession number SAMN12285942.

#### ACKNOWLEDGMENT

We would like to thank Dr. Lilian Praça for the technical assistance.

## REFERENCES

- 1 Ben-Dov E. 2014. *Bacillus thuringiensis* subsp. *israelensis* and its dipteran-specific toxins. *Toxins*. 6:1222–1243. doi: 10.3390/toxins6041222.
- 2 Martins ES, Praça LB, Dumas VF, Silva-Werneck JO, Sone EH, Waga IC, Berry C, Monnerat RG. 2007. Characterization of *Bacillus thuringiensis* isolates toxic to cotton boll weevil (*Anthonomus grandis*). *Biol Control*. 40:65–68. doi:10.1016/j.biocontrol.2006.09.009
- 3 - Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 28:1647-1649. doi: 10.1093/bioinformatics/bts199.
- 4 - Delcher AL, Kasif S, Fleischmann RD, Peterson J, White O, Salzberg SL. 1999. Alignment of whole genomes. *Nucleic Acids Res*. 27:2369–2376. doi: 10.1093/nar/27.11.2369.
- 5 – Schattner P, Brooks AN, Lowe TM. 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res*. 33:W686-W689. doi: 10.1093/nar/gki366.
- 6 - Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res*. 35:3100–3108. doi: 10.1093/nar/gkm160.
- 7 - Boeckmann B, Bairoch A, Apweiler R, Blatter MC, Estreicher A, Gasteiger E, Martin MJ, Michoud K, O’Donovan C, Phan I, Pilbout S, Schneider M. 2003. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res*. 31:365–370. [http:// dx.doi.org/10.1093/nar/gkg095](http://dx.doi.org/10.1093/nar/gkg095).
- 8 - Tatusov RL, Natale DA, Garkavtsev IV, Tatusova TA, Shankavaram UT, Rao BS, Kiryutin B, Galperin MY, Fedorova ND, Koonin EV. 2001. The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res*. 29:22–28. [http:// dx.doi.org/10.1093/nar/29.4.e22](http://dx.doi.org/10.1093/nar/29.4.e22).
- 9 - Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T, Yamanishi Y. 2008. KEGG for linking genomes to life and the environment. *Nucleic Acids Res*. 36:D480–D484. <http://dx.doi.org/10.1093/nar/gkm882>.
- 10 - 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. <http://dx.doi.org/10.1093/nar/gni118>.
- 11 – Crickmore N, Berry C, Panneerselvam S, Mishra R, Connor TR, Bonning BC. A structure-based nomenclature for *Bacillus thuringiensis* and other bacteria- derived

pesticidal proteins. Journal of Invertebrate Pathology.  
<https://doi.org/10.1016/j.jip.2020.107438>