Draft genome sequence of *Bacillus thuringiensis* strain S1307, an isolate toxic for lepidoptera

Projecto de sequência genómica de *Bacillus thuringiensis* strain S1307, um tóxico isolado para lepidoptera

DOI:10.34115/basrv6n3-011

Recebimento dos originais: 18/02/2022
Aceitação para publicação: 14/03/2022

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

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

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 Nacionale 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

The bacterium Bacillus thuringiensis (Bt) produces multiple toxins with activity against a diverse range of insect, including Diptera, Coleoptera and Lepidoptera. The Bt strain S1307 contains genes encoding Cry1Ab, Cry1Bc, Cry1Ia, Cry1J, Cry2Aa, Mpp5Aa, Spp1Aa, Vpb1B, Vpb2Bb and Vip3Aa. Genes coding for plant growth promotion

metabolic pathway were identified. And the operon which encodes the bacteriocins thuricin 17 and nisin were identified.

Keywords: sequencing, spodoptera frugiperda, bacteriocins.

RESUMO
A bactéria Bacillus thuringiensis (Bt) produz múltiplas toxinas com actividade contra uma gama diversificada de insectos, incluindo Diptera, Coleoptera e Lepidoptera. A estirpe Bt S1307 contém genes que codificam Cry1Ab, Cry1Bc, Cry1Ia, Cry1J, Cry2Aa, Mpp5Aa, Spp1Aa, Vpb1B, Vpb2Bb e Vip3Aa. Foram identificados os códigos Genes para a via metabólica de promoção do crescimento das plantas. E foram identificados o ópero que codifica as bacteriocinas thuricina 17 e nisina.

Palavras-chave: sequenciação, spodoptera frugiperda, bacteriocinas.

1 INTRODUCTION

Bacillus thuringiensis is an aerobic, rod-shaped, Gram-positive bacterium that produces crystal insect-toxic proteins during its sporulation process (1). These crystal toxins are solubilized and activated in the insect midgut. These toxins bind to specific receptors in the midgut cells causing cell lysis and insect death (2). This work describes the draft genome sequence of *B. thuringiensis* S1307.

Purified genomic and plasmid DNA`s from strain S1307 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 509 contigs totaling 6,385,657 bp (Q20 = 97,98%), with a maximum scaffold size of 249,195 bp, an N_{50} length of 59,619 bp, and 34.414% 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 6,407 coding sequences (4) and 34 tRNA (5) and 3 rRNA operons (6), and 5 ncRNA were predicted. The chromosome was sized in approx. 5,773,269 bp and the plasmids sizes ranged from 342,958 bp to 8,235 bp. 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 aminoacid 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. https://www.bpprc.org/). The strain S1307 draft genome sequence carries insecticidal toxin genes showing identities to the *Cry1Ab*, *Cry1Bc*, *Cry1Ia*, *Cry1J*, *Cry2Aa*, *Vpb1B*, *Vpb2Bb* and *Vip3Aa*. It was identified the genes which encodes the metabolic pathway of plant growth promotion 1-aminocyclopropane-1-carboxylate deaminase, indolepyruvate decarboxylase, putative acid phosphatase and siderophore biosynthesis protein. Using the new nomenclature of proteins toxic to insects proposed by Crickmore et al. (11) we found *Mpp5Aa* and *Spp1Aa*. The genomic analyses could identify the operon related to the antimicrobial peptide bacteriocin thuricin 17 and nisin. No parasporin gene was identified.

Nucleotide sequence accession numbers. The sequence of the *B. thuringiensis* strain S1307 has been deposited in GenBank with the accession number SAMN12286296.

ACKNOWLEDGMENT

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

REFERENCES

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

2 Castro BMC, Martinez LC, Barbosa SG, Serrão JE, Wilcken CF, Soares MA, Silva AA, Carvalho AG, Zanuncio JC. 2019. Toxicity and cytopathology mediated by *Bacillus thuringiensis* in the midgut of *Anticarsia gemmatalis* (Lepidoptera: Noctuidae). Scientific Reports. 9:6667. doi.org/10.1038/s41598-019-43074-0.

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. 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. 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