Draft genome sequence of *bacillus thuringiensis* S906, a toxic strain to coleoptera and lepidoptera orders

Projeto de sequência do genoma de *bacillus thuringiensis* S906, uma cepa tóxica para as ordens coleoptera e lepidoptera

DOI:10.34115/basrv6n3-006

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

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

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

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

Bacillus thuringiensis is an important bacterium which shows insecticide action against several pests. This work describes the draft genome sequence of B. thuringiensis S906, that showed toxicity against insects from Lepidoptera order (Anticarsia gemmatalis, Helicoverpa armigera, Plutella xylostella and Spodoptera frugiperda), and Coleoptera order (Anthonomus grandis) which contains the genes cry1Ba, Vpa, Vpb and Spp1A.

Genes coding for bacteriocin, plant growth promotion pathway, tyrosinase and enhancin were also identified.

Keywords: whole generation sequencing, biological control, cry toxin.

RESUMO
Bacillus thuringiensis é uma bactéria importante que mostra uma acção insecticida contra várias pragas. Este trabalho descreve o projecto de sequência genómica de B. thuringiensis S906, que mostrou toxicidade contra insectos da ordem Lepidoptera (Anticarsia gemmatalis, Helicoverpa armigera, Plutella xylostella e Spodoptera frugiperda), e da ordem Coleoptera (Anthonomus grandis) que contém os genes cry1Ba, Vpa, Vpb e Spp1A. Foram também identificados genes codificadores de bacteriocina, caminho de promoção do crescimento das plantas, tirosinase e enhancinase.

Palavras-chave: sequenciação de toda a geração, controlo biológico, toxina criogénica.

1 INTRODUCTION

Bacillus thuringiensis (Bt) is an ubiquitous Gram-positive, rod-shaped and sporulating bacterium that has been isolated worldwide from a great diversity of ecosystems (1). The Cry, Cyt and Vip proteins are considered the main components of the *B. thuringiensis* toxins which were encoded by genes that are usually located in plasmids and, less frequently, on the bacterial chromosome (2).

Embrapa Genetic Resources and Biotechnology has a collection of 3,000 strains of *Bacillus* spp., obtained from the soil from different regions of Brazil. Many of these strains are toxic to agricultural import insects. One of them, called S906, presented toxicity to insects of the Lepidoptera order (*Anticarsia gemmatalis*, *Helicoverpa armigera*, *Plutella xylostella* and *Spodoptera frugiperda*) and Coleoptera order (*Anthonomus grandis*).

Purified genomic and plasmid DNA`s from strain S906 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 245 contigs totaling 6,708,761 bp (Q20 = 99.22%), with a maximum scaffold size of 286,87 bp, an N_{50} length of 125,87 bp, and 33.89% 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 6,808 coding sequences (4), 85 tRNA (5), 3 rRNA operons (6) 5 ncRNA and 1 tmRNA (7) were predicted in the

genome. The chromosome was sized in approx. 5,334,534 bp and the plasmids sizes ranged from 585,136 bp to 2,038 bp. The functions of encoding genes were annotated by using the NCBI nr, Swiss-Prot (8), Clusters of Orthologous Groups (COG) (9), KEGG (10), and InterProScan (11) databases.

One Bt insecticidal toxin database was constructed with nucleotide and aminoacid sequences of *cry*, *vpa/vpb*, *spp*, *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 S906 draft genome sequence carries insecticidal toxin genes showing identities to the *cry1Ba1*, *vpa2Bb4*, *vpb1Bc1* and *spp1Aa1*, and *mpp5Aa1*. The genes which encodes the metabolic pathway of plant growth promotion 1-aminocyclopropane-1-carboxylate (ACC) deaminase, acid phosphatase, indole pyruvate decarboxylase (ipdC) and siderophore biosynthesis protein were identified. No parasporin gene was identified. The genomic analyses could identify the operon related to the antimicrobial peptide nisin and thuricin 4A, and one gene which encodes the tyrosinase but, this strain did not show any black colony during its growing in culture. One gene which encodes the enhancin protein was identified. The enhancin protein is considered as a metalloprotease showing a pathogenicity factor able to disrupt the protective peritrophic matrix (12). This strategy was useful to identify the *Vpa* and *Vpb* genes isolated in the genome once *vpa/vpb* genes can be identified in the *B. thuringiensis* genomes isolated or in operon. Using the new nomenclature of proteins toxic to insects proposed by Crickmore et al. (13) we found *mpp5A* and *spp1A*.

Nucleotide sequence accession numbers. The sequence of the *B. thuringiensis* strain S906 has been deposited in GenBank with the accession number SAMN12169346.

ACKNOWLEDGMENT

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

REFERENCES

1 - Palma L, Muñoz D, Berry C, Murillo J, Caballero P. 2014. *Bacillus thuringiensis* toxins: an overview of their biocidal activity. Toxins. 6:3296–3325. doi: 10.3390/toxins6123296

2 - Ibrahim MA, Griko N, Junker M, Bulla LA. 2010. Bacillus thuringiensis A genomics and proteomics perspective. Bioeng Bugs. Jan-Feb; 1(1): 31–50. doi: 10.4161/bbug.1.1.10519

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 – Panwar B, Arora A, Raghava GP. 2014. Prediction and classification of ncRNAs using structural information. BMC Genomics 15, 127. https://doi.org/10.1186/1471-2164-15-127

8 - 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

9 - 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

10 - 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.

11 - 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.

12 - Lepore L S, Roelvink P R, Granados R R. 1996. Enhancin, granulosis virus protein that facilitates nucleopolyhedrovirus (NPV) infections, is a metalloprotease. J Invertebr Pathol. 68:131–140. doi: 10.1006/jipa.1996.0070

13 – 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