Draft genome sequence of *Bacillus thuringiensis* strain S1905, an isolate toxic for lepidoptera

Projeto de sequência genómica de *Bacillus thuringiensis* strain S1905, um tóxico isolado para lepidoptera

DOI:10.34115/basrv6n3-012

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

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

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

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 Cry, Cyt, Vip and, Sip protein families show activity against insects of the orders Lepidoptera, Coleoptera and, Diptera. The Cry action toxins is to lyse midgut epithelial cells by forming pores in the target membrane. This work describes the draft genome sequence of B. thuringiensis S1905, which contains genes encoding Cry1A, Cry1I, Cry2A, Cry8K, and Vip3A. We found the spp1Aa gene and genes coding for plant

growth promotion metabolic pathway and Thuricin 17. One gene encoding the toxin enhancin and a locus which encodes the CRISPR were identified.

Keywords: plutella xylostella, ngs, shotgun.

RESUMO
As famílias Bacillus thuringiensis Cry, Cyt, Vip e, Sip protein mostram actividade contra os insectos das ordens Lepidoptera, Coleoptera e, Diptera. As toxinas Cry action toxins são as células epiteliais de acção lítica do meio do intestino, formando poros na membrana alvo. Este trabalho descreve o projecto da sequência genómica de B. thuringiensis S1905, que contém genes que codificam Cry1A, Cry1I, Cry2A, Cry8K, e Vip3A. Encontrámos o gene spp1Aa e genes que codificam a via metabólica de promoção do crescimento das plantas e Thuricin 17. Foram identificados um gene que codifica a toxina melhoradora e um locus que codifica o CRISPR.

Palavras-chave: plutella xylostella, ngs, shotgun.

1 INTRODUCTION

Bacillus thuringiensis (Bt) are spore-forming Gram-positive bacteria showing insecticide properties. Bt produce insecticidal crystal proteins during the sporulation phase and these crystals proteins (Cry) are called delta-endotoxins. Cry proteins are parasporal inclusion (Cry) proteins that exhibit toxic effect to insects (1). Different genomes of *B. thuringiensis* strains have been reported (2) and in this work we report the complete genome sequence of *B. thuringiensis* strain S1905.

Purified genomic and plasmid DNA`s from strain S1905 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 417 contigs totaling 6,367,384 bp (Q20 = 99.34%), with a maximum scaffold size of 222,970 bp, an N_{50} length of 53,586 bp, and 34.30% 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 6,274 coding sequences (4) and 56 tRNA (5) and 3 rRNA operons (6), and 5 ncRNA were predicted. The chromosome was sized in approx. 6,917,074 bp and the plasmids sizes ranged from 284,150 bp to 2,061 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 2195 draft genome sequence carries insecticidal toxin genes showing identities to the *Cry1A*, *Cry1I*, *Cry2A*, *Cry8K*, and *Vip3A*. 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 the *Spp1A* gene. The genomic analyses could identify the operon related to the antimicrobial peptide bacteriocin (thuricin 17). No parasporin gene was identified. In our genomic findings we detected a locus which encodes the CRISPR nuclease. This enzyme had much attention for its biotechnological potential.

The strain S1905 showed toxic activity against *S. frugiperda* LC_{50} (18 ng/cm^2), *A. gemmatalis* (LC_{50} 3.3 ng/cm^2) and *P. xylostella* (LC_{50} 1.46 ug/ml). These results were significatively when compared to other strains described in the work of Monnerat et al (12).

Nucleotide sequence accession numbers. The sequence of the *B. thuringiensis* strain S1905 has been deposited in GenBank with the accession number SAMN12289228.

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 – Sheppard AE, Poehlein A, Rosenstiel P, Liesegang H, Schulenburg H. 2013. Complete genome sequence of *Bacillus thuringiensis* strain 407 Cry-. Genome Announc. 1(1):e00158-12. doi:10.1128/genomeA.00158-12. doi:10.1128/genomeA.00158-12.

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.

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.

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

12 - Monnerat RG, Batista AC, Medeiros PT, Martins ES, Melatti VM, Praca LB, Dumas VF, Morinaga C, Demo C, Gomes ACM, Falcao R, Siqueira CB, Silva-Werneck JO, Berry C. 2007. Screening of Brazilian *Bacillus thuringiensis* isolates active against *Spodoptera frugiperda*, *Plutella xylostella* and *Anticarsia gemmatalis*. Biological Control 41:291–295. doi:10.1016/j.biocontrol.2006.11.008.