Draft genome sequence of *Bacillus thuringiensis* strain S601, a toxic strain to *Anthonomous grandis*

Sequência do genoma de *Bacillus thuringiensis* strain S601, uma estirpe tóxica para os *Anthonomous grandis*

DOI:10.34115/basrv6n3-005

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

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

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 showing insecticide proteins effective to control several agricultural pests, including boll weevil, Anthonomus grandis Boh. (Coleoptera: Curculionidae). This work describes the draft genome sequence of B. thuringiensis S601, which contains genes encoding the parasporal crystal cry1B, including other genes, like Vpa, Vpb, Sip and new ones described by new nomenclature

of Bt toxins. Genes coding for bacteriocin, plant growth promotion pathway and tyrosinase were also identified.

Keywords: next generation sequencing, biocontrol, boll weevil.

RESUMO
Bacillus thuringiensis é uma bactéria importante que mostra proteínas insecticidas eficazes para controlar várias pragas agrícolas, incluindo o gorgulho de Boll, Anthonomus grandis Boh. (Coleoptera: Curculionidae). Este trabalho descreve o projecto de sequência genómica de B. thuringiensis S601, que contém genes que codificam o cristal parasporal cry1B, incluindo outros genes, como Vpa, Vpb, Sip e novos, descritos pela nova nomenclatura de toxinas Bt. Os genes que codificam a bacteriocina, a via de promoção do crescimento das plantas e a tirosinase foram também identificados.

Palavras-chave: sequenciação da próxima geração, biocontrolo, gorgulho de boll.

1 INTRODUCTION

Bacillus thuringiensis (Bt) is a bacterium that produces parasporal protein crystals (Cry) toxic to some insect orders (1). Cry toxins exert their pathological effect by forming lytic pores in the membrane of insect midgut cells (2). *B. thuringiensis* has been used widely as a biopesticide, and there is a great interest to understand its pathogenic properties and how host evolve resistance. *B. thuringiensis* S601 was isolated from Brazil soil sample in our laboratory and is deposited in the Collection of Embrapa Genetic Resources and Biotechnology (3). This strain produce proteins 140 kDa and 65 kDa and showed a LC_{50} of 140 $mg.mL^{-1}$, lower than other *B. thuringiensis* strains like S1806, S1122 (*Bt* subsp. *tenebrionis*) and S1189 (*Bt* subsp. *israelensis*) that showed LC_{50} of 300 $mg.mL^{-1}$, 320 $mg.mL^{-1}$ and 740 $mg.mL^{-1}$ respectively (4).

Different genomes of *B. thuringiensis* strains have been published (5, 6) and in this work we report the complete genome sequence of *B. thuringiensis* strain S601.

Purified genomic and plasmid DNA's from strain S601 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 238 contigs totaling 6,871,004 bp (Q20 = 99.38%), with a maximum scaffold size of 424-501 bp, an N_{50} length of 146,010 bp, and 33.62% G+C content and genome coverage depth was approximately 100x. De novo assembly was carried out using the Geneious version (8.0.4) (7). A total of 6,863 coding sequences

(8) and 83 tRNA (9), 2 rRNA operons (10), 5 ncRNA and 1 tmRNA (11) were predicted in the genome. The chromosome was sized in approx. 5,297,893 bp and the plasmids sizes ranged from 597,988 bp to 2,023 bp. The functions of encoding genes were annotated by using the NCBI nr, Swiss-Prot (12), Clusters of Orthologous Groups (COG) (13), KEGG (14), and InterProScan (15) databases.

The custom insecticidal toxin database was constructed with nucleotide and amino acid sequences of *cry*, *vpa/vpb*, *sip*, *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 S601 draft genome sequence carries insecticidal toxin genes with identities to *cry1Ba1*, *Vpa2Bb4*, *Vpb1Bc1*, and *Sip1Aa1*. 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. The database analysis showed one antimicrobial peptide related to the bacteriocin (nisin and thuricin (A) and one tyrosinase gene. No parasporin gene was identified. This strategy was useful to identify the operon which encodes the *Vpa*/*Vpb* genes. Using the new nomenclature of proteins toxic to insects proposed by Crickmore et al. (16) we found *Mpp5Aa1* and *Spp1Aa1*. The strain S601 showed good results for controlling the cotton boll weevil *A. grandis*.

Nucleotide sequence accession numbers. The sequence of the *B. thuringiensis* strain S601 has been deposited in GenBank with the following accession number SAMN12046285.

ACKNOWLEDGMENT

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

REFERENCES

1 - Schnepf E, Crickmore N, Rie JV, Lereclus D, Baum J, Feitelson J, et al. 1998. *Bacillus thuringiensis* and its pesticidal crystal proteins. Microbiol Mol Biol Rev 62:775–806. PMCID: PMC98934.

2 - DeMaagd RA, Bravo A, Crickmore N. 2001. How *Bacillus thuringiensis* has evolved specific toxins to colonize the insect world. Trends Genet 17(4):193-199. http://dx.doi.org/10.1016/S0168-9525(01)02237-5.

3 - Monnerat RG, Silva SF, Silva-Werneck JO. 2001. Catalogo do Banco de germoplasma de bactérias entomopatogênicas do gênero *Bacillus*. Embrapa Recursos Genéticos e Biotecnologia, Brasília, 65 pp. ISSN:0102-0110.

4 - 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. https://doi.org/10.1016/j.biocontrol.2006.09.009.

5 - 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.

6 - He J, Shao X, Zheng H, Li M, Wang J, Zhang Q, Li L, Liu Z, Sun M, Wang S, Yu Z. 2010. Complete genome sequence of *Bacillus thuringiensis* mutant strain BMB171. J Bacteriol 192:4074– 4075. doi: 10.1128/JB.00562-10.

7 - 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.

8 - 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.

9 – 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.

10 - 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.

11 - 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.

12 - 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.

13 - 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.

14 - 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.

15 - 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.

16 - Crickmore N, Berry C, Panneerselvam S, Mishra R, Connor TR, Bonning BC. 2020. A structure-based nomenclature for *Bacillus thuringiensis* and other bacteriaderived pesticidal proteins. Journal of Invertebrate Pathology. https://doi.org/10.1016/j.jip.2020.107438