Microbial Pathogenesis 108 (2017) 55–60

Contents lists available at ScienceDirect

Microbial Pathogenesis

journal homepage: www.elsevier.com/locate/micpath

Complete genome sequence and bioinformatics analyses of *Bacillus thuringiensis* strain BM-BT15426

Junyan Liu [a, b], Lin Li [a, c], Brian M. Peters [b], Bing Li [a, c], Dingqiang Chen, Ph.D [d, **], Zhenbo Xu, Ph.D [a, c, e, *], Mark E. Shirtliff [e]

[a] School of Food Science and Engineering, South China University of Technology, Guangzhou 510640, PR China
[b] Department of Clinical Pharmacy, College of Pharmacy, University of Tennessee Health Science Center, Memphis TN 38163, USA
[c] Guangdong Province Key Laboratory for Green Processing of Natural Products and Product Safety, Guangzhou 510640, PR China
[d] Department of Laboratory Medicine, First Affiliated Hospital of Guangzhou Medical University, Guangzhou 510120, PR China
[e] Department of Microbial Pathogenesis, School of Dentistry, University of Maryland, Baltimore MD 21201, USA

ARTICLE INFO

Article history:
Received 13 April 2017
Received in revised form
27 April 2017
Accepted 2 May 2017
Available online 4 May 2017

Keywords:
Bacillus thuringiensis
Genetic characteristics
Complete genome
Bioinformatics analyses

ABSTRACT

Objectives: This study aimed to investigate the genetic characteristics of *Bacillus thuringiensis* strain BM-BT15426.
Methods: *B. thuringiensis* strain was identified by sequencing the PCR product (amplifying *16S rRNA* gene) using ABI Prism 377 DNA Sequencer. The genome was sequenced using PacBio RS II sequencers and assembled *de novo* using HGAP. Also, further genome annotation was performed.
Results: The genome of *B. thuringiensis* strain BM-BT15426 has a length of 5,246,329 bp and contains 5409 predicted genes with an average G + C content of 35.40%. Three genes were involved in the "Infectious diseases: Amoebiasis" pathway. A total of 21 virulence factors and 9 antibiotic resistant genes were identified.
Conclusions: The major pathogenic factors of *B. thuringiensis* strain BM-BT15426 were identified through complete genome sequencing and bioinformatics analyses which contributes to further study on pathogenic mechanism and phenotype of *B. thuringiensis*.

1. Introduction

Bacillus thuringiensis was initially discovered as a pathogen of various insects because of the production of insecticidal toxin. It is frequently used as an insecticidal agent due to the specific toxin activity towards insects and the lack of toxicity to animals. Some toxin genes specific for a variety of insects have been studied and used in genetically modified plants which have been engineered to produce the toxin themselves. *B. thuringiensis* has also been proposed as an alternative treatment method of several parasites (helminths, ticks, mites) that infect domestic animals. The synthesis of protein origin crystalline inclusions that are highly toxic to arthropods of Lepidoptera, Diptera and Coleoptera has been characterized for *B. thuringiensis* [1,2]. The toxicity of *B. thuringiensis* on mites has also been studied in ectoparasites such as Tetranychus and synanthropic mites [3,4]. *In vitro*, spores and crystals of *B. thuringiensis* was reported to induce over 90% mortality rate of Varroa destructor mites, which have affection on bees [5]. It eliminated the importance of *B. thuringiensis* spores and crystals which has also been verified in genetically modified plants.

The genome is the basis to thorough comprehension of the internal characteristic of a bacterium and the whole genome sequencing is an initial strategy to investigate the phenotypic behavior [6–9]. Thus, in this study, complete genome sequencing and bioinformatics analyses were conducted to investigate the genetic characteristics of *B. thuringiensis* strain BM-BT15426.

2. Material and methods

2.1. Bacterial strain

B. thuringiensis strain BM-BT15426 was isolated from the culture of *Lactobacillus* strains by accident and showed high survival rate

* Corresponding author. School of Food Science and Engineering, South China University of Technology, Guangzhou 510640, PR China.
** Corresponding author.
E-mail addresses: jyksys@126.com (D. Chen), zhenbo.xu@hotmail.com (Z. Xu).

http://dx.doi.org/10.1016/j.micpath.2017.05.006
0882-4010/

even in extreme conditions (Low pH, low oxygen and nutrients, etc.) [10,11]. It was grown in LB broth at 37 °C with shaking (200 rpm). Experiments were performed in a Biosafety Level 2 (BSL-2) biological safety cabinet.

2.2. Strain identification

To identify the bacteria, PCR was performed to amplify the bacterial *16S rRNA* gene using the forward (5′-AGAGTTTGATCCTGGCTCAG-3′) and reverse primers (5′-CTACGGCTACCTTGTTACGA-3′) [12,13]. The total DNA was extracted from the bacterial strain using a bacterial DNA extraction kit (Sigma-Aldrich, USA). The PCR reactions were carried out in a EDC-810 Thermo cycler (Eastwin Life Sciences, Inc., China), and the cycling profile consisted of an initial heating at 94 °C for 3min, followed by 30 cycles of 94 °C for 50 s, 50 °C for 50 s, and 72 °C for 1 min, with a final 10 min extension at 72 °C. Following amplification, the PCR product was sequenced with ABI Prism 377 DNA Sequencer (PE Applied Biosystems, USA) and compared with published *16S rRNA* gene sequences in a NCBI BLAST program [14–16].

2.3. Genome sequencing and assembly

Total DNA was isolated from a culture of *B. thuringiensis* strain BM-BT15426 by using a bacterial DNA extraction kit (Sigma-Aldrich, USA) and quantified by using pulsed field gel electrophoresis (PFGE) and Qubit dsDNA BR assay (Life Technologies, USA) [17]. Genome DNA was fragmented using Covaris® g-TUBE® Shearing Device and then purified with 0.45× AMPure beads. Then the fragmented DNA underwent DNA repair and end repair. After purified with 0.45× AMPure beads, blunt hairpin adapters were ligated into blunt-end of DNAs [18]. ExoIII and ExoVII enzymes were used to digest DNA not ligated to adapters. Finally, the desired insert size of SMRTbellTM library was selected using BluePippin. The quantified SMRTbellTM libraries were then sequenced using PacBio RS II sequencers. The sequences were filtered and assembled *de novo* using HGAP [19].

2.4. Bioinformatics analyses

Genes in the genome of *B. thuringiensis* strain BM-BT15426 were predicted using GeneMarkS [20] and further annotated by Nr (https://www.ncbi.nlm.nih.gov/protein/), SwissProt (http://www.gpmaw.com/html/swiss-prot.html), COG [21], KEGG [22], and GO [23] databases. *tRNA*, *rRNA*, and repeat sequences were predicted by tRNAscan-SE (v1.21) [24], RNAmmer (v1.2) [25], and RepeatMasker (http://www.repeatmasker.org/), respectively. Importantly, antibiotic resistance genes and virulence factors were predicted by Antibiotic Resistance Genes Database (ARDB, http://www.mgc.ac.cn/VFs/main.htm) and Virulence Factors of Pathogenic Bacteria Database (VFDB, http://ardb.cbcb.umd.edu/), respectively. Also, protein domains (Pfam), pathogen host interactions (PHI), carbohydrate-active enzymes (CAZy), prophages, and secretions were predicted by Pfam_Scan (ftp://ftp.sanger.ac.uk/pub/databases/Pfam/Tools/), PHI-base [26], CAZy database [27], PHAST 2.1 [28], and siganlP 4.1 [29], respectively.

2.5. GenBank accession number

The complete genome sequence of *B. thuringiensis* strain BM-BT15426 has been submitted to GenBank under accession number CP020723.

3. Results and discussion

3.1. Bacterial identification

The *16S rRNA* gene has been amplified (Fig. 1) and sequenced to perform bacterial identification. The *16S rRNA* gene sequence of the strain showed 99.9% similarity with that of *B. thuringiensis* strain BAB-Bt2 (GenBank accession no. AM293345). Thus, the bacterial strain was identified as *B. thuringiensis*.

3.2. General features of B. thuringiensis genome

The genome of *B. thuringiensis* strain BM-BT15426 was assembled to one scaffold. It has a length of 5,246,329 bp (Fig. 2) and contains 5409 predicted genes with an average G + C content of 35.40%. The genome shows high similarity (90% coverage and 99% identity) with the genome of *B. thuringiensis* strain HD682 (GenBank accession number: CP009720). However, the genome of *B. thuringiensis* strain BM-BT15426 was 45 kb shorter than that of *B. thuringiensis* strain HD682 [30]. It indicated some sequences missing during the evolution of *B. thuringiensis* strain BM-BT15426 [31,32]. A total of 74 repeat sequences, 5S *rRNA*, 16S *rRNA*, 23S *rRNA* and 106 *tRNA* were identified in the genome. There were 5110 and 3556 predicted genes have COG function annotation (Fig. 3) and GO function annotation (Fig. 4), respectively. Also, 2626 predicted genes were involved in 82 KEGG pathways (Table S1) which were classified into “Metabolism” (66/82), “Genetic Information Processing” (13/82), “Cellular Processes” (1/82), “Human Diseases” (1/82), and “Environmental Information Processing” (1/82). Three genes including *rocF* encoded arginase, *ahpC* encoded peroxiredoxin, and an unnamed gene encoded aldehyde dehydrogenase family protein were involved in the “Infectious diseases: Amoebiasis” pathway which was the only pathway related to “Human

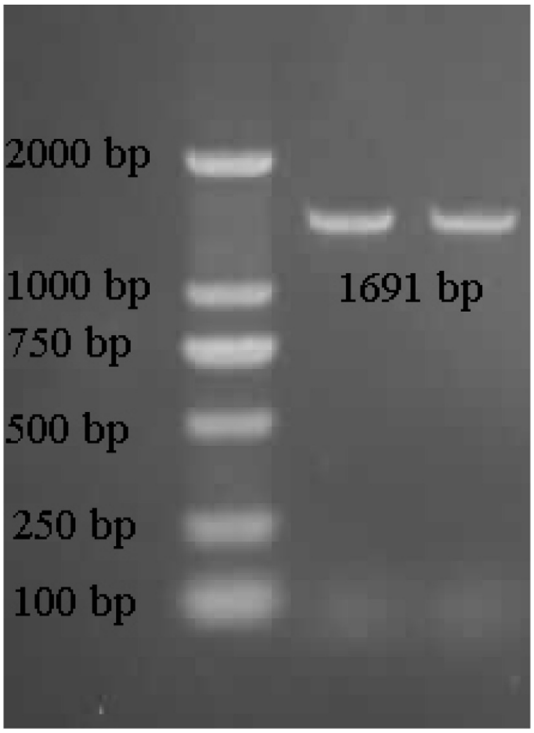

Fig. 1. PCR amplification of *16S rRNA* gene from *B. thuringiensis* strain BM-BT15426. The left lane is marker (DSTM2000), the middle and right lanes are amplification bands of *16S rRNA* gene from *B. thuringiensis* strain BM-BT15426.