



# Draft Genome Sequence of *Bacillus* sp. Strain IITD106, Isolated from Oil-Contaminated Soil

Arif Nissar Zargar,<sup>a</sup> Saroj Mishra,<sup>a</sup> Manoj Kumar,<sup>b</sup>  Preeti Srivastava<sup>a</sup>

<sup>a</sup>Department of Biochemical Engineering and Biotechnology, Indian Institute of Technology Delhi, New Delhi, India

<sup>b</sup>Indian Oil Corporation, R&D Centre, Faridabad, India

**ABSTRACT** Here, we report the whole-genome sequence of *Bacillus* sp. strain IITD106. The bacterium has the unique ability to produce saponins. The complete nucleotide sequence will provide insights into the various genes and regulators involved in the biosynthesis of saponin.

**B** *acillus* sp. strain IITD106 was isolated from crude oil-contaminated soil by enrichment culture using asphaltene as a carbon source (1). The bacterium was found to produce a unique biosurfactant, saponin, which is usually reported from plants (2). Whole-genome sequencing was essential not only to understand the different genes involved in the biosynthesis of saponin but also to explore other catabolic abilities of the bacterium.

For extraction of the genomic DNA, the culture was inoculated into Luria broth and incubated at 30°C for 24 h. Genomic DNA was isolated using the DNeasy PowerSoil Pro kit (catalog number 47016; Qiagen). To determine the whole-genome sequence, a library was prepared using the NEBNext Ultra II FS DNA library preparation kit with sample purification beads (catalog number E6177; New England Biolabs). Briefly, the genomic DNA was enzymatically fragmented by targeting 200–300 bp fragments. End repairing of the fragments was performed to create blunt ends, and adenylation was performed by adding a single A to the 3' ends. This was followed by ligation with loop adapters and cleavage by the uracil-specific excision reagent provided in the kit. The DNA was purified using AMPure beads and enriched using six cycles with NEBNext Ultra II Q5 master mix, Illumina universal primers, and sample-specific octamer primers. The sequences of the universal Illumina flow cell adapters were as follows: Illumina P5 adaptor, 5'-AATGATACGGCGACCACCGAGATCTACAC-3'; Illumina P7 adaptor, 5'-CAAGCAGAAGACGGCATACGAGAT-3'. The bar code indexes of the octamers were as follows: P7 index, UDI16, AACGCATT; P5 index, UDI16, CAACCTGC.

Paired-end sequencing was carried out using the Illumina HiSeq 4000 platform. Trim Galore v0.6.6 was used with default settings for trimming and removal of the adapter sequences (3). The read length was 150 bp, and raw reads were assembled using Unicycler v0.4.8 with default settings (4). The total number of reads obtained was 7,800,000. The total size of the assembly obtained was 4,486,399 bp, distributed in 143 contigs, and the GC content was 41%. The  $N_{50}$  value for the assembled contigs was 82,061 bp, and the sequence coverage was 269×.

The genome was annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v5.2 (5). A total of 4,385 protein-coding genes and 90 pseudogenes were predicted. The BLAST search results revealed 99.72% identity to *Bacillus wuyishanensis* (6).

The gene annotation analysis of *Bacillus* sp. strain IITD106 revealed the presence of several genes involved in the biosynthesis of saponin. The antibiotic resistance genes were identified *in silico* using CARD (7). Interestingly, genes for resistance to antibiotics such as doxorubicin, tetracycline, linearmycin, fosmidomycin, bicyclomycin, and tunicamycin were detected. A cluster of *ssuACB* genes, which are responsible for the specific uptake of aliphatic sulfonates, was identified. Genes involved in polyaromatic hydrocarbon degradation,

**Editor** Steven R. Gill, University of Rochester School of Medicine and Dentistry

**Copyright** © 2021 Zargar et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Preeti Srivastava, preeti@dbeb.iitd.ac.in.

**Received** 3 October 2021

**Accepted** 31 October 2021

**Published** 2 December 2021

such as catechol 2,3-dioxygenase and the putative ring cleavage extradiol dioxygenase MhqA, were identified in the genome of this organism.

**Data availability.** The whole-genome sequence of *Bacillus* sp. strain IITD106 has been deposited in DDBJ/ENA/GenBank under accession number [JAIAZY000000000](https://doi.org/10.1186/s13568-021-01285-7), BioProject number [PRJNA750262](https://doi.org/10.1186/s13568-021-01285-7), BioSample number [SAMN20447436](https://doi.org/10.1186/s13568-021-01285-7), and SRA accession number [SRX12347337](https://doi.org/10.1186/s13568-021-01285-7).

## ACKNOWLEDGMENT

We thank the Department of Biotechnology, Government of India, for financial support.

## REFERENCES

1. Zargar AN, Kumar A, Sinha A, Kumar M, Skiadas I, Mishra S, Srivastava P. 2021. Asphaltene biotransformation for heavy oil upgradation. *AMB Express* 11:127. <https://doi.org/10.1186/s13568-021-01285-7>.
2. Zargar AN, Lympertou A, Skiadas I, Kumar M, Srivastava P. 2022. Structural and functional characterization of a novel biosurfactant from *Bacillus* sp. IITD106. *J Hazard Mater* 423:127201. <https://doi.org/10.1016/j.jhazmat.2021.127201>.
3. Krueger F. 2012. Trim Galore: a wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files, with some extra functionality for MspI-digested RRBS-type (Reduced Representation Bisulfite-Seq) libraries. [https://www.bioinformatics.babraham.ac.uk/projects/trim\\_galore](https://www.bioinformatics.babraham.ac.uk/projects/trim_galore).
4. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>.
5. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
6. Liu B, Liu G-H, Sengonca C, Schumann P, Che J-M, Zhu Y-J, Wang J-P. 2015. *Bacillus wuyishanensis* sp. nov., isolated from rhizosphere soil of a medical plant, *Prunella vulgaris*. *Int J Syst Evol Microbiol* 65:2030–2035. <https://doi.org/10.1099/ijs.0.000215>.
7. McArthur AG, Wagglechner N, Nizam F, Yan A, Azad MA, Baylay AJ, Bhullar K, Canova MJ, De Pascale G, Ejim L, Kalan L, King AM, Koteva K, Morar M, Mulvey MR, O'Brien JS, Pawlowski AC, Piddock LJV, Spanogiannopoulos P, Sutherland AD, Tang I, Taylor PL, Thaker M, Wang W, Yan M, Yu T, Wright GD. 2013. The Comprehensive Antibiotic Resistance Database. *Antimicrob Agents Chemother* 57:3348–3357. <https://doi.org/10.1128/AAC.00419-13>.