

# Draft Genome Sequence of *Anoxybacillus suryakundensis* Strain JS1<sup>T</sup> (DSM 27374<sup>T</sup>) Isolated from a Hot Spring in Jharkhand, India

Kamal Deep,<sup>a</sup> Abhijit Poddar,<sup>a</sup> William B. Whitman,<sup>b</sup> Subrata K. Das<sup>a</sup>

Department of Biotechnology, Institute of Life Sciences, Nalco Square, Bhubaneswar, India<sup>a</sup>; Department of Microbiology, University of Georgia, Athens, Georgia, USA<sup>b</sup>

***Anoxybacillus suryakundensis* strain JS1<sup>T</sup>, a facultative anaerobic, moderately thermophilic, alkalitolerant bacterium, was isolated from a hot spring. The estimated genome is 2.6 Mb and encodes 2,668 proteins.**

Received 17 June 2016 Accepted 17 June 2016 Published 11 August 2016

**Citation** Deep K, Poddar A, Whitman WB, Das SK. 2016. Draft genome sequence of *Anoxybacillus suryakundensis* strain JS1<sup>T</sup> (DSM 27374<sup>T</sup>) isolated from a hot spring in Jharkhand, India. *Genome Announc* 4(4):e00824-16. doi:10.1128/genomeA.00824-16.

**Copyright** © 2016 Deep 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 Subrata K. Das, [subratkdas@hotmail.com](mailto:subratkdas@hotmail.com).

*Anoxybacillus suryakundensis* strain JS1<sup>T</sup> is a Gram-positive, nonmotile, facultative anaerobe. A moderately thermophilic and alkalitolerant bacterium, it was isolated from sediment samples from a hot spring in Suryakund, Jharkhand, India. Heterotrophic growth occurred at 40 to 60°C and pH 5.5 to 11.5 (1).

The draft genome of *A. suryakundensis* was sequenced using the Illumina HiSeq 2000 platform, which generated 11,125,280 reads totaling 1,679.9 Mb. Filtered Illumina reads were assembled using Velvet version 1.2.07 (2), and 1- to 3-kb simulated paired-end reads were created from Velvet contigs using wgsim version 0.3.0 (<https://github.com/lh3/wgsim>). Illumina reads were assembled with simulated read pairs using Allpaths-LG version r46652 (3).

The genome was annotated using the JGI Microbial Genome Annotation Pipeline (4). Genes were identified using Prodigal (5), followed by manual curation using GenePRIMP (6) for finished genomes and draft genomes in fewer than 20 scaffolds (>50 kb) containing 92.9% of the genome. The final assembly was based on 1,510.0 Mb of Illumina data with 302-fold input read coverage. The predicted coding sequences (CDSs) were translated and used to search the NCBI nonredundant database and the UniProt, TIGRFam, Pfam, KEGG, COG, and InterPro databases. The tRNAscanSE tool (7) was used to find tRNA genes, whereas rRNA genes were found by searches against models of the rRNA genes built from SILVA (8). Other noncoding RNAs, such as the RNA components of the protein secretion complex and the RNase P, were identified by searching the genome for the corresponding Rfam profiles using Infernal (<http://infernal.janelia.org>).

The final draft assembly contained 44 contigs in 41 scaffolds, totaling 2.6 Mb with an *N*<sub>50</sub> contig size of 148.6 kb. The largest contig was 245.9 kb with a G+C content of 42.0 mol%. The draft genome sequence has 2,668 CDSs, 61 tRNAs, 16 rRNAs, and two clustered regularly interspaced short palindromic repeats.

**Accession number(s).** This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number [LIOK00000000](https://www.ncbi.nlm.nih.gov/nuclink/LIOK00000000). The version described in this paper is version LIOK01000000.

## ACKNOWLEDGMENTS

This work was supported in part by the funding received by S.K.D. from the Department of Biotechnology, Government of India (D.O. no. BT/

PR9712/NBD/52/91/2007). K.D. and A.P. acknowledge the University Grant Commission (UGC), Government of India, and Council of Scientific and Industrial Research (CSIR), Government of India, New Delhi, for providing the research fellowship.

We also acknowledge the generous assistance of Nikos C. Kyrpides, Tanja Woyke, Nicole Shapiro, and the other members of the JGI microbial genome sequencing team. The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy under contract no. DE-AC02-05CH11231.

## REFERENCES

- Deep K, Poddar A, Das SK. 2013. *Anoxybacillus suryakundensis* sp. nov., a moderately thermophilic, alkalitolerant bacterium isolated from hot spring at Jharkhand, India. *PLoS One* 8:e85493. <http://dx.doi.org/10.1371/journal.pone.0085493>.
- Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18:821–829. <http://dx.doi.org/10.1101/gr.074492.107>.
- Gnerre S, MacCallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, Sharpe T, Hall G, Shea TP, Sykes S, Berlin AM, Aird D, Costello M, Daza R, Williams L, Nicol R, Gnirke A, Nusbaum C, Lander ES, Jaffe DB. 2011. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc Natl Acad Sci U S A* 108:1513–1518. <http://dx.doi.org/10.1073/pnas.1017351108>.
- Mavromatis K, Ivanova NN, Chen IM, Szeto E, Markowitz VM, Kyrpides NC. 2009. The DOE-JGI standard operating procedure for the annotations of microbial genomes. *Stand Genomic Sci* 1:63–67. <http://dx.doi.org/10.4056/sigs.632>.
- Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. <http://dx.doi.org/10.1186/1471-2105-11-119>.
- Pati A, Ivanova NN, Mikhailova N, Ovchinnikova G, Hooper SD, Lykidis A, Kyrpides NC. 2010. GenePRIMP: a gene prediction improvement pipeline for prokaryotic genomes. *Nat Methods* 7:455–457. <http://dx.doi.org/10.1038/nmeth.1457>.
- Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25:955–964. <http://dx.doi.org/10.1093/nar/25.5.0955>.
- Pruesse E, Quast C, Knittel K, Fuchs B, Ludwig W, Peplies J, Glöckner FO. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* 35:2188–2196. <http://dx.doi.org/10.1093/nar/gkm864>.