Draft Genome Sequence of *Pannonibacter indicus* Strain HT23^T (DSM 23407^T), a Highly Arsenate-Tolerant Bacterium Isolated from a Hot Spring in India

Saumya Bandyopadhyay,[a] **William B. Whitman,**[b] **Subrata K. Das**[a]

Department of Biotechnology, Institute of Life Sciences, Bhubaneswar, India[a]; Department of Microbiology, University of Georgia, Athens, Georgia, USA[b]

ABSTRACT *Pannonibacter indicus* strain HT23^T, a highly arsenate-tolerant bacterium, was isolated from a tropical hot spring. The estimated genome is 4.2 Mb with 3,818 protein-coding sequences containing putative genes, some of which are involved in arsenate resistance.

*P*annonibacter indicus* strain HT23^T is a motile, aerobic, Gram-negative, and rod-shaped bacterium belonging to the class *Alphaproteobacteria*. This bacterium was isolated from a tropical hot spring at Athamallik, Orissa, India (1, 2). Strain HT23^T is able to grow in the presence of 500 mM of sodium arsenate in low-phosphate medium. The genome sequence will help to understand the mechanism of arsenic tolerance in this bacterium.

The draft genome of strain HT23^T was generated at the DOE Joint Genome Institute (JGI). An Illumina standard shotgun library was constructed and sequenced using the Illumina HiSeq 2000 platform (3), which generated 10,175,130 paired-end reads totaling 1,536.4 Mb. Filtered Illumina reads were assembled using Velvet (4), wgsim, and Allpaths-LG (5). The final draft assembly contained 38 contigs in 36 scaffolds, totaling 4.2 Mb, with an input read coverage of 261.8-fold. The largest and N_{50} contigs were 646.4 kb and 216.1 kb, respectively, with a G+C content of 63.5%.

The genome was annotated using the JGI Microbial Annotation Pipeline (6). Genes were identified using Prodigal (7), followed by manual curation using GenePRIMP (8). The tRNAs, rRNAs, and other noncoding RNA genes were identified by searching the genome using the tRNAScan-SE tool (9), rRNA gene models built from SILVA (10), and Infernal (https://www.janelia.org/publication/infernal-10-inference-rna-alignments), respectively. The draft genome sequence has 3,818 candidate protein-coding sequences (CDSs), 50 tRNAs, 7 rRNAs, 13 other RNAs, and 3 clustered regularly interspaced short palindromic repeats (CRISPRs).

Accession number(s). This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number LIPT00000000. The version described in this paper is the first version, LIPT01000000.

REFERENCES

1. Bandyopadhyay S, Schumann P, Das SK. 2013. *Pannonibacter indica* sp. nov., a highly arsenate-tolerant bacterium isolated from a hot spring in India. Arch Microbiol 195:1–8. https://doi.org/10.1007/s00203-012-0840-z.

2. Bandyopadhyay S, Schumann P, Das SK. 2015. List of new names and new combinations previously effectively, but not validly, published-validation List no. 150. Int J Syst Evol Microbiol 63:797–798. https://doi.org/10.1099/ijsem.0.000464.

3. Bennett S. 2004. Solexa Ltd. Pharmacogenomics 5:433–438. https://doi.org/10.1517/14622416.5.4.433.

4. Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 18:821–829. https://doi.org/10.1101/gr.074492.107.

5. 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. https://doi.org/10.1073/pnas.1017351108.

6. 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. https://doi.org/10.4056/sigs.632.

7. 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. https://doi.org/10.1186/1471-2105-11-119.

8. 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. https://doi.org/10.1038/nmeth.1457.

9. 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. https://doi.org/10.1093/nar/25.5.0955.

10. Pruesse E, Quast C, Knittel K, Fuchs B, Ludwig W, Peplies J, Glckner FO. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res 35:2188–7196. https://doi.org/10.1093/nar/gkt1209.