

Draft Genome Sequence of *Chelatococcus sambhunathii* Strain HT4^T (DSM 18167^T) Isolated from a Hot Spring in India

Jhasketan Badhai,^a William B. Whitman,^b Subrata K. Das^a

Department of Biotechnology, Institute of Life Sciences, Nalco Square, Bhubaneswar, India^a; Department of Microbiology, University of Georgia, Athens, Georgia, USA^b

The moderately thermophilic bacterium *Chelatococcus sambhunathii* strain HT4^T was isolated from hot spring sediment. Based upon the draft genome sequence, the genome is 4.4 Mb and encodes 4,147 proteins.

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

Citation Badhai J, Whitman WB, Das SK. 2016. Draft genome sequence of *Chelatococcus sambhunathii* strain HT4^T (DSM 18167^T) isolated from a hot spring in India. *Genome Announc* 4(4):e00825-16. doi:10.1128/genomeA.00825-16.

Copyright © 2016 Badhai 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.

Chelatococcus sambhunathii strain HT4^T is a Gram-negative, aerobic rod. It is nonsporulating and motile, belonging to the class *Alphaproteobacteria*. It was isolated from a hot spring located in Athamallik, Orissa, India (1). It grows heterotrophically at pH 6.0 to 8.5 and at 20 to 50°C.

The draft genome of *C. sambhunathii* strain HT4^T was generated at the DOE Joint Genome Institute (JGI) using the Illumina HiSeq 2000 platform (2). An Illumina standard shotgun library was constructed and sequenced using the Illumina HiSeq 2000 platform, which generated 8,684,234 reads totaling 1,311.3 Mb. The filtered Illumina reads were assembled using Velvet (3), wgsim (<https://github.com/lh3/wgsim>), and Allpaths-LG (4) tools. The final draft assembly contains 38 contigs in 35 scaffolds, totaling 4.4 Mb, with an input read coverage of 193.7-fold. The largest and *N*₅₀ contigs are 742.1 kb and 426.1 kb, respectively, with a G+C content of 67.8%.

The genome was annotated using the JGI Microbial Genome Annotation Pipeline (5). Genes were identified using the Prodigal (6) and GenePRIMP (7) programs. The tRNAs, rRNAs, and other noncoding RNA genes were identified by searching the genome using the tRNAscan-SE tool (8), rRNA gene models built from SILVA (9), and Infernal (<http://infernal.janelia.org>), respectively. The draft genome contains 4,147 coding sequences, 49 tRNAs, 4 rRNAs, 13 other RNAs, 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 [LIOL00000000](https://www.ncbi.nlm.nih.gov/nuclink/LIOL00000000). The version described in this paper is version LIOL01000000.

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). J.B. acknowledges the Department of Biotechnology, 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

- Panday D, Das SK. 2010. *Chelatococcus sambhunathii* sp. nov., a moderately thermophilic alphaproteobacterium isolated from hot spring sediment. *Int J Syst Evol Microbiol* 60:861–865. <http://dx.doi.org/10.1099/ijs.0.013466-0>.
- Bennett S. 2004. Solexa Ltd. *Pharmacogenomics* 5:433–438. <http://dx.doi.org/10.1517/14622416.5.4.433>.
- 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/signs.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* 35:7188–7196. <http://dx.doi.org/10.1093/nar/gkm864>.