

Genome Sequence of Strain HIMB30, a Novel Member of the Marine *Gammaproteobacteria*

Megan J. Huggett and Michael S. Rappé

Hawaii Institute of Marine Biology, School of Ocean and Earth Science and Technology, University of Hawaii at Manoa, Kaneohe, Hawaii, USA

Strain HIMB30 was isolated from coastal Hawaii seawater by extinction culturing in seawater-based oligotrophic medium. It is a phylogenetically unique member of the class *Gammaproteobacteria* that is only distantly related to its closest cultured relatives. Here we present the genome sequence of strain HIMB30, including genes for proteorhodopsin-based phototrophy and the Calvin-Benson-Bassham cycle.

Strain HIMB30 was isolated via dilution-to-extinction culturing methods (3) from surface seawater collected off the coast of Oahu, HI, in the subtropical Pacific Ocean. It is of interest because it belongs to a novel, previously uncultivated lineage of *Gammaproteobacteria* that is found in cultivation-independent surveys of marine bacterioplankton, including coastal picoplankton communities from the continental shelf of North Carolina (11), the coast of Sapelo Island in Georgia (10), and Plum Island Sound in Massachusetts (1). The two most closely related described species, *Litoricola lipolytica* and *L. marina* in the family *Litoricolaceae* (2, 6), share just 89.9% 16S rRNA gene sequence similarity with strain HIMB30. Given the unique phylogenetic status of strain HIMB30, it most likely represents a novel family within the order *Oceanospirillales*.

The genome of HIMB30 was shotgun sequenced by the J. Craig Venter Institute as part of the Moore Foundation Microbial Genome Sequencing Project (<http://camera.calit2.net/microgenome>), resulting in 629,573 reads of 454 FLX Titanium sequence data. A draft, unclosed genome was assembled to 102× coverage from 624,954 reads (254,771,663 bases) of sequence data. Functional annotation was performed with the Integrated Microbial Genomes Expert Review (IMG-ER) pipeline (9). Genes identified using Prodigal (5) were translated and used to search the NCBI nonredundant protein database and the UniProt, TIGRFam, Pfam, PRIAM, KEGG, COG, and InterPro databases. Nontranslated genes were predicted using tRNAscanSE (8), RNAmmer (7), and Rfam (4). Additional gene prediction analysis and manual functional annotation were performed within IMG-ER.

The draft genome of HIMB30 consists of 36 contigs ranging in size from 266 to 638,152 bases, resulting in a total genome of 2,168,870 bases. The G+C content is 49.86%. In all, 2,295 predicted open reading frames and 2,249 predicted protein-coding genes were identified, of which 1,874 have a predicted function. There are predicted single copies of the 5S, 16S, and 23S rRNA genes and 36 predicted tRNAs. Genes for a complete tricarboxylic acid cycle are present in the genome of HIMB30, but it possesses lesions in both the Embden-Meyerhoff-Parnas pathway and the oxidative portion of the pentose phosphate pathway. Genes for two key enzymes of the Entner-Doudoroff pathway, phosphogluconate dehydratase and 2-keto-3-deoxyphosphogluconate aldolase, were found. The genome contains putative genes for the synthesis of all of the essential amino acids and a number of vitamins, including

para-aminobenzoic acid, lipoic acid, pyridoxine (vitamin B₆), riboflavin, pantothenic acid, and coenzyme A. In addition to genes for carbon monoxide and sulfur oxidation, genes for proteorhodopsin-based phototrophy and CO₂ fixation via the Calvin-Benson-Bassham cycle were found.

Nucleotide sequence accession numbers. This Whole Genome Shotgun project has been deposited in GenBank under accession no. [AGIG0000000](https://doi.org/10.1093/genbank/AGIG0000000). The version described in this paper is the first version, AGIG01000000. The genome project was also deposited in the Genomes OnLine Database as project Gi03571, and the NCBI taxonomy identification number is 751994.

ACKNOWLEDGEMENTS

We thank Steve Ferriera, Justin Johnson, and other J. Craig Venter Institute scientists for performing the sequencing and assembly of this genome, with support provided by the Gordon and Betty Moore Foundation Microbial Genome Sequencing Project (<http://www.moore.org/marinemicro>). Strain preparation, genomic DNA extraction, and annotation were supported by the Center for Microbial Oceanography: Research and Education (NSF Science and Technology Center award EF-0424599).

REFERENCES

- Acinas SG, et al. 2004. Fine-scale phylogenetic architecture of a complex bacterial community. *Nature* 430:551–554.
- Choi A, Lee K, Oh H-M, Feng J, Cho J-C. 2010. *Litoricola marina* sp. nov. *Int. J. Syst. Evol. Microbiol.* 60:1303–1306.
- Connon SA, Giovannoni SJ. 2002. High-throughput methods for culturing microorganisms in very-low-nutrient media yield diverse new marine isolates. *Appl. Environ. Microbiol.* 68:3878–3885.
- Griffiths-Jones S, et al. 2005. Rfam: annotating non-coding RNAs in complete genomes. *Nucleic Acids Res.* 33:D121–D124.
- Hyatt D, et al. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119.
- Kim H, Choo Y-J, Cho J-C. 2007. *Litoricolaceae* fam. nov., to include *Litoricola lipolytica* gen. nov., sp. nov., a marine bacterium belonging to the order *Oceanospirillales*. *Int. J. Syst. Evol. Microbiol.* 57:1793–1798.
- Lagesen K, et al. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35:3100–3108.
- Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detec-

Received 10 November 2011 Accepted 11 November 2011

Address correspondence to Michael S. Rappé, rappem@hawaii.edu.

Copyright © 2012, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JB.06506-11

- tion of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25: 955–964.
9. Markowitz VM, Ivanova NN, Chen IMA, Chu K, Kyrpides NC. 2009. IMG ER: a system for microbial genome annotation expert review and curation. *Bioinformatics* 25:2271–2278.
 10. Poretsky RS, Sun S, Mou X, Moran MA. 2010. Transporter genes expressed by coastal bacterioplankton in response to dissolved organic carbon. *Environ. Microbiol.* 12:616–627.
 11. Rappé MS, Kemp PF, Giovannoni SJ. 1997. Phylogenetic diversity of marine coastal picoplankton 16S rRNA genes cloned from the continental shelf off Cape Hatteras, North Carolina. *Limnol. Oceanogr.* 42:811–826.