




Draft Genome Sequence of the Predatory Marine Bacterium *Halobacteriovorax* sp. Strain JY17

Jonathan M. Young,^a  Timofey Skvortsov,^a Ksenia Arkhipova,^{a*} Christopher C. R. Allen^a

^aSchool of the Biological Sciences, Queen's University Belfast, Belfast, United Kingdom

ABSTRACT A draft genome sequence of *Halobacteriovorax* sp. strain JY17 was assembled from a metagenomic data set. The 3.47-Mbp genome of this unusual predatory bacterium contains 3,263 protein-coding sequences, 33 tRNAs, and 2 copies each of the 16S, 23S, and 5S rRNA genes. This is only the third sequenced representative of this genus.

The genus *Halobacteriovorax* comprises predatory deltaproteobacteria found in marine environments (1). They prey on Gram-negative bacteria, penetrating the cell wall and multiplying within the periplasmic space (2, 3), and have been identified as potential biocontrol agents (4). Here, we present *Halobacteriovorax* sp. strain JY17, assembled from a methanol-enriched seawater sample from the Irish Sea. To our knowledge, this is only the third available genome for a member of the genus *Halobacteriovorax* to date.

Methanol enrichment cultures were set up according to a previously described method (5) and incubated in a shaking incubator for 18 days at 22°C and 50 rpm, after which DNA was extracted from cell pellets using the PowerSoil DNA isolation kit (MO BIO Laboratories, Carlsbad, CA, USA). Metagenomic DNA libraries were prepared using an Illumina Nextera DNA library kit and sequenced on an Illumina NextSeq 500 DNA sequencer (Illumina, San Diego, CA, USA) in mid-output mode using a paired-end flow cell (2 × 150 bp read length, V2 chemistry) at the DNA Sequencing Facility, Department of Biochemistry, University of Cambridge (Cambridge, UK).

Raw reads from three individual samples were pre-processed using BBTools (6) and coassembled using metaSPAdes (7). Contigs were taxonomically assigned using Kaiju (8) and binned on the basis of GC% and differential coverage using CONCOCT (9).

The resulting bins were manually refined to reduce contamination using Anvi'o (10); briefly, for each bin, contigs were clustered on the basis of GC% by Euclidean distance and Ward's linkage clustering followed by visualization of the resulting dendrograms and manual removal of any outliers. Reads mapping to contigs within each genome bin were extracted and reassembled using SPAdes (11) with the “-careful” flag enabled. A final scaffolding step was performed using SSPACE (12). Final draft genome completeness and contamination was assessed using CheckM (13). The final draft genome was annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (14) and reviewed using RAST v2.0 (15).

This resulted in a 3.47-Mbp draft genome estimated to be 93.9% complete with 1.79% potential contamination, comprising 5 scaffolds with a GC content of 36%, 3,354 coding sequences, 34 tRNAs, and 2 copies of 16S, 23S, and 5S rRNA genes. Interestingly, only 26% of the coding sequences could be assigned to SEED subsystems, indicating an unusually high level of novel functional diversity.

BLAST alignment of the 16S rRNA genes indicated 98.7% similarity to *Halobacteriovorax marinus*; however, further investigation with gANI (16) revealed a 77.6% average nucleotide identity and an alignment fraction of 0.76 between the two genomes. Additionally, species assignment using spec1 (17) revealed an average identity of 86.5%

Received 13 November 2017 **Accepted** 14 November 2017 **Published** 4 January 2018

Citation Young JM, Skvortsov T, Arkhipova K, Allen CCR. 2018. Draft genome sequence of the predatory marine bacterium *Halobacteriovorax* sp. strain JY17. Genome Announc 6:e01416-17. <https://doi.org/10.1128/genomeA.01416-17>.

Copyright © 2018 Young 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 Christopher C. R. Allen, c.allen@qub.ac.uk.

* Present address: Ksenia Arkhipova, Theoretical Biology and Bioinformatics, Utrecht University, Utrecht, The Netherlands.

to *H. marinus* across 40 conserved housekeeping genes. Therefore, we propose that the newly obtained genome represents a new species within the genus *Halobacteriovorax*. Furthermore, our data indicate that alignment of the 16S rRNA gene may be a poor method of delineating species level diversity in the genus *Halobacteriovorax*.

Accession number(s). The metagenome-derived draft genome sequence of *Halobacteriovorax* sp. strain JY17 has been deposited in DDBJ/EMBL/GenBank under the accession number [NJER00000000](#).

ACKNOWLEDGMENT

This work was partly funded by the Agri-Food Quest Competence Centre (AFQCC) at Queen's University Belfast, Belfast, United Kingdom.

REFERENCES

- Koval SF, Williams HN, Stine OC. 2015. Reclassification of *Bacteriovorax marinus* as *Halobacteriovorax marinus* gen. nov., comb. nov. and *Bacteriovorax litoralis* as *Halobacteriovorax litoralis* comb. nov.; description of *Halobacteriovoraceae* fam. nov. in the class *deltaproteobacteria*. *Int J Syst Evol Microbiol* 65:593–597. <https://doi.org/10.1099/ijs.0.070201-0>.
- Williams HN, Lymperopoulou DS, Athar R, Chauhan A, Dickerson TL, Chen H, Laws E, Berhane TK, Flowers AR, Bradley N, Young S, Blackwood D, Murray J, Mustapha O, Blackwell C, Tung Y, Noble RT. 2016. *Halobacteriovorax*, an underestimated predator on bacteria: potential impact relative to viruses on bacterial mortality. *ISME J* 10:491–499. <https://doi.org/10.1038/ismej.2015.129>.
- Chen H, Young S, Berhane TK, Williams HN. 2012. Predatory *Bacteriovorax* communities ordered by various prey species. *PLoS One* 7:e34174. <https://doi.org/10.1371/journal.pone.0034174>.
- Cao H, Hou S, He S, Lu L, Yang X. 2014. Identification of a *Bacteriovorax* sp. isolate as a potential biocontrol bacterium against snakehead fish-pathogenic *Aeromonas veronii*. *J Fish Dis* 37:283–289. <https://doi.org/10.1111/jfd.12120>.
- Janvier M, Frehel C, Grimont F, Gasser F. 1985. *Methylophaga marina* gen. nov., sp. nov. and *Methylophaga thalassica* sp. nov., marine methylotrophs. *Int J Syst Bacteriol* 35:131–139. <https://doi.org/10.1099/0020713-35-2-131>.
- Bushnell B. 2014. BMAP short read aligner, and other bioinformatic tools. <https://sourceforge.net/projects/bbmap>.
- Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. *Genome Res* 27:824–834. <https://doi.org/10.1101/gr.213959.116>.
- Menzel P, Ng KL, Krogh A. 2016. Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat Commun* 7:11257. <https://doi.org/10.1038/ncomms11257>.
- Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, Lahti L, Loman NJ, Andersson AF, Quince C. 2014. Binning metagenomic contigs by coverage and composition. *Nat Methods* 11:1144–1146. <https://doi.org/10.1038/nmeth.3103>.
- Eren AM, Esen ÖC, Quince C, Vineis JH, Morrison HG, Sogin ML, Delmont TO. 2015. Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* 3:e1319. <https://doi.org/10.7717/peerj.1319>.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
- Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27:578–579. <https://doi.org/10.1093/bioinformatics/btq683>.
- Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>.
- 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>.
- Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 9:75. <https://doi.org/10.1186/1471-2164-9-75>.
- Varghese NJ, Mukherjee S, Ivanova N, Konstantinidis KT, Mavrommatis K, Kyrpides NC, Pati A. 2015. Microbial species delineation using whole genome sequences. *Nucleic Acids Res* 43:6761–6771. <https://doi.org/10.1093/nar/gkv657>.
- Mende DR, Sunagawa S, Zeller G, Bork P. 2013. Accurate and universal delineation of prokaryotic species. *Nat Methods* 10:881–884. <https://doi.org/10.1038/nmeth.2575>.