



High-Quality Draft Genome Sequences of the Uncultured Delta3 Endosymbiont (*Deltaproteobacteria*) Assembled from Metagenomes of the Gutless Marine Worm *Olavius algarvensis*

Yui Sato,^a Juliane Wippler,^a Cecilia Wentrup,^b Tanja Woyke,^c Nicole Dubilier,^{a,d} Manuel Kleiner^e

^aMax Planck Institute for Marine Microbiology, Bremen, Germany

^bUniversity of Vienna, Department of Microbiology and Ecosystem Science, Vienna, Austria

^cDOE Joint Genome Institute, Berkeley, California, USA

^dUniversity of Bremen, Faculty 2 Biology/Chemistry, Bremen, Germany

^eNorth Carolina State University, Department of Plant and Microbial Biology, Raleigh, North Carolina, USA

ABSTRACT Here, we present two high-quality, draft metagenome-assembled genomes of deltaproteobacterial OalgDelta3 endosymbionts from the gutless marine worm *Olavius algarvensis*. Their 16S rRNA gene sequences share 98% identity with Delta3 endosymbionts of related host species *Olavius ilvae* (GenBank accession no. [AJ620501](#)) and *Inanidrilus exumae* (GenBank accession no. [FM202060](#)), for which no symbiont genomes are available.

Gutless phalloporine annelids are a species-rich, monophyletic group of marine oligochaetes that lack a digestive system (1). Instead, they rely on phylogenetically diverse bacterial symbionts that provide them with nutrition via chemosynthesis (2–5). Symbionts of different host species show host specificity but often form closely related bacterial clades. To date, none of the symbionts has been cultivated, but comparative genomic analyses have been performed using high-throughput metagenomic sequencing (3, 6). Here, we present two draft genomes of a deltaproteobacterial symbiont from the host *Olavius algarvensis*, termed OalgDelta3. OalgDelta3 is closely related to uncultured Delta3 endosymbionts from two other gutless phalloporinids (6, 7), namely, *Olavius ilvae* and *Inanidrilus exumae*, sharing 98% sequence identity with their 16S rRNA genes (GenBank accession no. [AJ620501](#) and [FM202060](#), respectively), and to deltaproteobacteria (SEEP-SRB1) that form consortia with anaerobic methane oxidizers in anoxic marine sediments (8, 9).

Two *O. algarvensis* specimens were collected from Elba, Italy (specimen B2SA from Sant'Andrea [42°48'26.00"N, 10°8'28.00"E] in April 2013 and specimen A1CA from Cavoli [42°44'4.15"N, 10°11'10.47"E] in June 2014), as described previously (4), and stored at 4°C in RNAlater (Ambion Life Technologies). DNA was extracted from individual specimens using the Qiagen blood and tissue kit. Two separate DNA libraries were prepared from mechanically sheared (Covaris LE220), size-selected (using Beckman Coulter solid-phase reversible immobilization beads targeting 300 bp) total DNA using the KAPA library creation kit for Illumina platforms (Kapa Biosystems) and underwent paired-end sequencing on the Illumina HiSeq 2000 platform (2 × 150 bp) at the DOE Joint Genome Institute. A total of 53,614,071 and 36,269,211 raw read pairs were obtained from B2SA and A1CA, respectively.

Reads were quality controlled with BBDuk (minimum kmer, 11; minimum length, 36 bp; minimum Phred quality score, 2) of the BBMap toolkit v36.86 (<https://sourceforge.net/projects/bbmap>) and BayesHammer implemented in SPAdes v3.9.1 (10, 11), and 53,474,255 (B2SA) and 36,195,816 (A1CA) clean read pairs were separately assembled with MEGAHIT v1.0.6 (12). Metagenome-assembled genomes (MAGs) were

Citation Sato Y, Wippler J, Wentrup C, Woyke T, Dubilier N, Kleiner M. 2020. High-quality draft genome sequences of the uncultured Delta3 endosymbiont (*Deltaproteobacteria*) assembled from metagenomes of the gutless marine worm *Olavius algarvensis*. *Microbiol Resour Announc* 9:e00704-20. <https://doi.org/10.1128/MRA.00704-20>.

Editor Frank J. Stewart, Georgia Institute of Technology

Copyright © 2020 Sato et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](#).

Address correspondence to Yui Sato, ysato@mpi-bremen.de.

Received 16 June 2020

Accepted 7 July 2020

Published 30 July 2020

obtained using MetaBAT v0.26.3 (13) and refined using Bandage v0.08.1 (14). Genome completeness and contamination were estimated using CheckM v1.0.5 (15). Average nucleotide identity (ANI) was calculated with enveomics tools (16). Genome annotation and analysis were performed with RASTtk (17–19) and Pathway Tools v21.0 (20). Default parameters were used for all software unless otherwise noted.

MAGs were assigned as OalgDelta3 by identifying the previously reported 16S rRNA gene sequence (GenBank accession no. [AM493254](#)) with >99.5% identity. The OalgDelta3 MAG from B2SA has 5.58 Mbp (401 contigs, with an N_{50} value of 76,958 bp and the largest contig of 424,721 bp; GC content, 54.2%; genome coverage, 50×) and contains 5,795 protein-coding sequences (CDSs) and 43 tRNA-coding regions. The OalgDelta3 MAG from A1CA has 5.74 Mbp (350 contigs, with an N_{50} value of 89,163 bp and the largest contig of 348,870 bp; GC content, 54.1%; genome coverage, 350×) and includes 5,876 CDSs and 45 tRNA-coding regions. The two MAGs share 99.86% identical 16S rRNA gene sequences, with a genome-wide ANI of 99.87%. Both MAGs have 94.4% estimated completeness, with no (B2SA) or very low (A1CA; 0.03%) contamination estimates, and conform to the MIMAG standards for high-quality draft genomes (21). Similar to other deltaproteobacterial symbionts of *O. algarvensis*, OalgDelta3 harbors pathways for sulfate reduction, carbon monoxide and H₂ oxidation, utilization of sugars, short-chain fatty acids and oligopeptides, and a multitude of associated ATP-binding cassette (ABC) transporters, tripartite ATP-independent periplasmic (TRAP)-type transporters, and other transporters.

Data availability. The MAGs of the OalgDelta3 endosymbiont have been deposited in the European Nucleotide Archive under accession no. [PRJEB28157](#) (assemblies [GCA_903231395](#) and [GCA_903231505](#)), using the data brokerage service of the German Federation for Biological Data (GFBio) (22), in compliance with the Minimal Information about any (x) Sequence (MIxS) standard (23). The raw sequences are available at the Sequence Read Archive (SRA) under accession no. [SRX2712534](#) and [SRX2554373](#).

ACKNOWLEDGMENTS

This work was supported by the Max Planck Society, a Moore Foundation Marine Microbial Initiative Investigator Award to N.D. (grant GBMF3811), the North Carolina State Chancellor's Faculty Excellence Program Cluster on Microbiomes and Complex Microbial Communities (M.K.), the USDA National Institute of Food and Agriculture Hatch Project 1014212 (M.K.), and the European Union's Horizon 2020 research and innovation program under Marie Skłodowska-Curie grant agreement 660280 (C.W.). Sequencing was conducted by the U.S. DOE Joint Genome Institute, a DOE Office of Science User Facility, and was supported under contract DE-AC02-05CH11231.

REFERENCES

- Dubilier N, Blazejak A, Ruehland C. (2005) Symbioses between bacteria and gutless marine oligochaetes, p 251–275. In Overmann J (ed), Molecular basis of symbiosis. progress in molecular and subcellular biology, vol 41. Springer, Berlin, Germany.
- Dubilier N, Mülders C, Ferdelman T, de Beer D, Perntaler A, Klein M, Wagner M, Erséus C, Thiermann F, Krieger J, Giere O, Amann R. 2001. Endosymbiotic sulphate-reducing and sulphide-oxidizing bacteria in an oligochaete worm. *Nature* 411:298–302. <https://doi.org/10.1038/35077067>.
- Woyke T, Teeling H, Ivanova NN, Huntemann M, Richter M, Gloeckner FO, Boffelli D, Anderson IJ, Barry KW, Shapiro HJ, Szeto E, Kyrpidis NC, Mussmann M, Amann R, Bergin C, Ruehland C, Rubin EM, Dubilier N. 2006. Symbiosis insights through metagenomic analysis of a microbial consortium. *Nature* 443:950–955. <https://doi.org/10.1038/nature05192>.
- Kleiner M, Wentrup C, Lott C, Teeling H, Wetzel S, Young J, Chang YJ, Shah M, VerBerkmoes NC, Zarzycki J, Fuchs G, Markert S, Hempel K, Voigt B, Becher D, Liebeke M, Lalk M, Albrecht D, Hecker M, Schweder T, Dubilier N. 2012. Metaproteomics of a gutless marine worm and its symbiotic microbial community reveal unusual pathways for carbon and energy use. *Proc Natl Acad Sci U S A* 109:E1173–E1182. <https://doi.org/10.1073/pnas.1121198109>.
- Wippler J, Kleiner M, Lott C, Gruhl A, Abraham PE, Giannone RJ, Young JC, Hettich RL, Dubilier N. 2016. Transcriptomic and proteomic insights into innate immunity and adaptations to a symbiotic lifestyle in the gutless marine worm *Olavius algarvensis*. *BMC Genomics* 17:942. <https://doi.org/10.1186/s12864-016-3293-y>.
- Kleiner M, Woyke T, Ruehland C, Dubilier N. 2011. The *Olavius algarvensis* metagenome revisited: lessons learned from the analysis of the low-diversity microbial consortium of a gutless marine worm, p 319–333. In de Bruijn FJ (ed), Handbook of molecular microbial ecology II: metagenomics in different habitats. John Wiley & Sons, Hoboken, NJ.
- Bergin C. 2009. Phylogenetic diversity and metabolic versatility of the bacterial endosymbionts in marine gutless oligochaete worms. Ph.D. thesis. University of Bremen, Bremen, Germany.
- Schreiber L, Holler T, Knittel K, Meyerdieks A, Amann R. 2010. Identification of the dominant sulfate-reducing bacterial partner of anaerobic methanotrophs of the ANME-2 clade. *Environ Microbiol* 12:2327–2340. <https://doi.org/10.1111/j.1462-2920.2010.02275.x>.
- Kleindienst S, Ramette A, Amann R, Knittel K. 2012. Distribution and *in situ* abundance of sulfate-reducing bacteria in diverse marine hydrocarbon seep sediments. *Environ Microbiol* 14:2689–2710. <https://doi.org/10.1111/j.1462-2920.2012.02832.x>.

10. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski 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>.
11. Nikolenko SI, Korobeynikov AI, Alekseyev MA. 2013. BayesHammer: Bayesian clustering for error correction in single-cell sequencing. *BMC Genomics* 14:S7. <https://doi.org/10.1186/1471-2164-14-S1-S7>.
12. Li D, Luo R, Liu CM, Leung CM, Ting HF, Sadakane K, Yamashita H, Lam TW. 2016. MEGAHIT v1.0: a fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods* 102:3–11. <https://doi.org/10.1016/j.ymeth.2016.02.020>.
13. Kang DD, Froula J, Egan R, Wang Z. 2015. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* 3:e1165. <https://doi.org/10.7717/peerj.1165>.
14. Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of *de novo* genome assemblies. *Bioinformatics* 31:3350–3352. <https://doi.org/10.1093/bioinformatics/btv383>.
15. 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>.
16. Rodriguez-R LM, Konstantinidis KT. 2016. The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. *PeerJ Prepr* 4:e1900v1. <https://doi.org/10.7287/peerj.preprints.1900v1>.
17. 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>.
18. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res* 42:D206–D214. <https://doi.org/10.1093/nar/gkt1226>.
19. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA, Stevens R, Vonstein V, Wattam AR, Xia F. 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep* 5:8365. <https://doi.org/10.1038/srep08365>.
20. Karp PD, Latendresse M, Paley SM, Krummenacker M, Ong QD, Billington R, Kothari A, Weaver D, Lee T, Subhraveti P, Spaulding A, Fulcher C, Keseler IM, Caspi R. 2016. Pathway Tools version 19.0 update: software for pathway/genome informatics and systems biology. *Brief Bioinform* 17:877–890. <https://doi.org/10.1093/bib/bbv079>.
21. Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, Schulz F, Jarett J, Rivers AR, Eloie-Fadrosch EA, Tringe SG, Ivanova NN, Copeland A, Clum A, Becraft ED, Malmstrom RR, Birren B, Podar M, Bork P, Weinstock GM, Garrity GM, Dodsworth JA, Yooshef S, Sutton G, Glöckner FO, Gilbert JA, Nelson WC, Hallam SJ, Jungbluth SP, Ettema TJG, Tighe S, Konstantinidis KT, Liu WT, Baker BJ, Rattei T, Eisen JA, Hedlund B, McMahon KD, Fierer N, Knight R, Finn R, Cochrane G, Karsch-Mizrachi I, Tyson GW, Rinke C, Lapidus A, Meyer F, Yilmaz P, Parks DH, Eren AM, Schriml L, Banfield JF, Hugenholtz P, Woyke T, Genome Standards Consortium. 2017. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat Biotechnol* 35:725–731. <https://doi.org/10.1038/nbt.3893>.
22. Diepenbroek M, Glöckner F, Grobe P, Güntsch A, Huber R, König-Ries B, Kostadinov I, Nieschulze J, Seeger B, Tolksdorf R, Triebel D. 2014. Towards an integrated biodiversity and ecological research data management and archiving platform: the German Federation for the Curation of Biological Data (GFBio), p 1711–1724. *In* Plödereder E, Grunskel L, Schneider E, Ull D (ed), *Informatik 2014: Big Data Komplexität meistern*. GI ed.: Lecture Notes in Informatics (LNI) Proceedings 232. Köllen Verlag, Bonn, Germany.
23. Yilmaz P, Kottmann R, Field D, Knight R, Cole JR, Amaral-Zettler L, Gilbert JA, Karsch-Mizrachi I, Johnston A, Cochrane G, Vaughan R, Hunter C, Park J, Morrison N, Rocca-Serra P, Sterk P, Arumugam M, Bailey M, Baumgartner L, Birren BW, Blaser MJ, Bonazzi V, Booth T, Bork P, Bushman FD, Buttigieg PL, Chain PS, Charlson E, Costello EK, Huot-Creasy H, Dawyndt P, DeSantis T, Fierer N, Fuhrman JA, Gallery RE, Gevers D, Gibbs RA, San Gil I, Gonzalez A, Gordon JI, Guralnick R, Hankeln W, Highlander S, Hugenholtz P, Jansson J, Kau AL, Kelley ST, Kennedy J, Knights D, Koren O, Kuczynski J, Kyrpides N, Larsen R, Lauber CL, Legg T, Ley RE, Lozupone CA, Ludwig W, Lyons D, Maguire E, Methé BA, Meyer F, Muegge B, Nakielny S, Nelson KE, Nemergut D, Neufeld JD, Newbold LK, Oliver AE, Pace NR, Palanisamy G, Peplies J, Petrosino J, Proctor L, Pruesse E, Quast C, Raes J, Ratnasingham S, Ravel J, Relman DA, Assunta-Sansone S, Schloss PD, Schriml L, Sinha R, Smith MI, Sodergren E, Spo A, Stombaugh J, Tiedje JM, Ward DV, Weinstock GM, Wendel D, White O, Whiteley A, Wilke A, Wortman JR, Yatsunenko T, Glöckner FO. 2011. Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIXS) specifications. *Nat Biotechnol* 29:415–420. <https://doi.org/10.1038/nbt.1823>.