



Heriot-Watt University  
Research Gateway

## Genome Sequence of *Roseovarius* sp. Strain MCTG156(2b) Isolated from a Phytoplankton Net Trawl on the Scottish West Coast

### Citation for published version:

Gutierrez, T, Whitman, WB, Huntemann, M, Copeland, A, Chen, A, Vargese, N, Kyrpides, NC, Pillay, M, Ivanova, N, Mikhailova, N, Mukherjee, S, Stamatis, D, Reddy, TBK, Ngan, CY, Chovatia, M, Daum, C, Shapiro, N & Woyke, T 2017, 'Genome Sequence of *Roseovarius* sp. Strain MCTG156(2b) Isolated from a Phytoplankton Net Trawl on the Scottish West Coast', *Genome Announcements*, vol. 5, no. 35, e00837-17. <https://doi.org/10.1128/genomeA.00837-17>

### Digital Object Identifier (DOI):

[10.1128/genomeA.00837-17](https://doi.org/10.1128/genomeA.00837-17)

### Link:

[Link to publication record in Heriot-Watt Research Portal](#)

### Document Version:

Publisher's PDF, also known as Version of record

### Published In:

Genome Announcements

### General rights

Copyright for the publications made accessible via Heriot-Watt Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

### Take down policy

Heriot-Watt University has made every reasonable effort to ensure that the content in Heriot-Watt Research Portal complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [open.access@hw.ac.uk](mailto:open.access@hw.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# Genome Sequence of *Roseovarius* sp. Strain MCTG156(2b) Isolated from a Phytoplankton Net Trawl on the Scottish West Coast

Tony Gutierrez,<sup>a</sup> William B. Whitman,<sup>b</sup> Marcel Huntemann,<sup>c</sup> Alex Copeland,<sup>c</sup> Amy Chen,<sup>c</sup> Neha Vargese,<sup>c</sup> Nikos C. Kyrpides,<sup>c</sup> Manoj Pillay,<sup>c</sup> Natalia Ivanova,<sup>c</sup> Natalia Mikhailova,<sup>c</sup> Supratim Mukherjee,<sup>c</sup> Dimitrios Stamatis,<sup>c</sup> T. B. K. Reddy,<sup>c</sup> Chew Yee Ngan,<sup>c</sup> Mansi Chovatia,<sup>c</sup> Chris Daum,<sup>c</sup> Nicole Shapiro,<sup>c</sup> Tanja Woyke<sup>c</sup>

School of Engineering and Physical Sciences, Heriot Watt University, Edinburgh, United Kingdom<sup>a</sup>;  
Department of Microbiology, University of Georgia, Athens, Georgia, USA<sup>b</sup>; DOE Joint Genome Institute,  
Walnut Creek, California, USA<sup>c</sup>

**ABSTRACT** *Roseovarius* sp. strain MCTG156(2b) was isolated from a phytoplankton net sample collected on the west coast of Scotland and was selected based on its ability to degrade polycyclic aromatic hydrocarbons. Here, we present the genome sequence of this strain, which is 5,113,782 bp, with 5,142 genes and an average G+C content of 60.7%.

*Roseovarius* sp. strain MCTG156(2b) was isolated from a phytoplankton net sample that was trawled in 2009 at a sampling station designated LY1, located on the west coast of Scotland near Oban, Argyll. The strain was isolated by enrichment with phenanthrene in Zobell's 2216 marine medium at a 10-fold dilution. Colonies on agar plates sprayed with phenanthrene produced distinct halos that indicated the strain's ability to degrade the hydrocarbon. Based on 16S rRNA gene sequence identity, the closest type species was *Roseovarius mucosus* strain DFL-24, which had also been isolated from a laboratory culture of a dinoflagellate (1).

Here, we report the genome sequence of *Roseovarius* sp. strain MCTG156(2b). Genomic DNA was sequenced through the DOE Joint Genome Institute 2014 Genomic encyclopedia of type strains phase III study (2) using Pacific Biosciences (PacBio) technology. A PacBio SMRTbell library was constructed and sequenced on the PacBio RS platform, which generated 170,293 filtered subreads totaling 598.2 Mbp. All general aspects of library construction and sequencing performed at the JGI can be found at <http://www.jgi.doe.gov>. The raw reads were assembled using HGAP (version 2.1.1) (3). The final draft assembly produced 8 scaffolds containing 8 contigs totaling 5.1 Mbp in size, with input read coverage of 121.5×

Project information is available in the Genomes OnLine Database (4). Genes were identified using Prodigal (5), as part of the Joint Genome Institute (JGI) Microbial Annotation Pipeline (6). The predicted coding sequences (CDSs) were translated and used to search the National Center for Biotechnology Information (NCBI) nonredundant, UniProt, TIGRFam, Pfam, KEGG, COG, and InterPro databases. The tRNAscan-SE tool (7) was used to find tRNA genes, whereas ribosomal RNA genes were found by searches against models of the ribosomal RNA genes built from SILVA (8). Other noncoding RNAs, such as the RNA components of the protein secretion complex and RNase P, were identified by searching the genome for the corresponding Rfam profiles using Infernal (<http://infernal.janelia.org>). Additional analysis and manual functional annotation were performed within the Integrated Microbial Genomes-Expert Review

Received 5 July 2017 Accepted 7 July 2017 Published 31 August 2017

**Citation** Gutierrez T, Whitman WB, Huntemann M, Copeland A, Chen A, Vargese N, Kyrpides NC, Pillay M, Ivanova N, Mikhailova N, Mukherjee S, Stamatis D, Reddy TBK, Ngan CY, Chovatia M, Daum C, Shapiro N, Woyke T. 2017. Genome sequence of *Roseovarius* sp. strain MCTG156(2b) isolated from a phytoplankton net trawl on the Scottish west coast. *Genome Announc* 5:e00837-17. <https://doi.org/10.1128/genomeA.00837-17>.

This is a work of the U.S. Government and is not subject to copyright protection in the United States. Foreign copyrights may apply.

Address correspondence to Tony Gutierrez, [tony.gutierrez@hw.ac.uk](mailto:tony.gutierrez@hw.ac.uk).

(IMG-ER) platform (<https://img.jgi.doe.gov/>) developed by the Joint Genome Institute, Walnut Creek, CA, USA (9).

The complete genome sequence length was 5,113,782 bp, with a G+C content of 60.7%. The genome contained 5,142 genes (5,078 protein-coding genes), with functional predictions for 4,198 of them. A total of 64 RNA genes were detected. Other genes characteristic for the genus are given in the IMG database (10).

**Accession number(s).** The draft genome sequence of *Roseovarius* sp. strain MCTG156(2b) obtained in this study was deposited in GenBank as part of BioProject no. PRJNA224116, with individual genome sequences submitted as whole-genome shotgun projects under the accession no. [JQLS00000000](https://doi.org/10.1093/jqs.0.63832-0).

## ACKNOWLEDGMENTS

T.G. was supported by a Marie Curie International Outgoing Fellowship (PIOF-GA-2008-220129) within the Seventh European Community Framework Program. Partial support was also provided by the European Union's Horizon 2020 research and innovation program under grant agreement no. 635340. The work was conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, which is supported by the Office of Science of the U.S. Department of Energy under contract no. DE-AC02-05CH11231.

## REFERENCES

1. Biebl H, Allgaier M, Lünsdorf H, Pukall R, Tindall BJ, Wagner-Döbler I. 2005. *Roseovarius mucosus* sp. nov., a member of the *Roseobacter* clade with trace amounts of bacteriochlorophyll *a*. *Int J Syst Evol Microbiol* 55:2377–2383. <https://doi.org/10.1099/ijs.0.63832-0>.
2. Whitman WB, Woyke T, Klenk HP, Zhou Y, Lilburn TG, Beck BJ, De Vos P, Vandamme P, Eisen JA, Garrity G, Hugenholtz P, Kyrpides NC. 2015. Genomic encyclopedia of bacterial and archaeal type strains, phase III: the genomes of soil and plant-associated and newly described type strains. *Stand Genomic Sci* 10:26. <https://doi.org/10.1186/s40793-015-0017-x>.
3. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Non-hybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <https://doi.org/10.1038/nmeth.2474>.
4. Mukherjee S, Stamatis D, Bertsch J, Ovchinnikova G, Verezhemskaya O, Isbandi M, Thomas AD, Ali R, Sharma K, Kyrpides NC, Reddy TB. 2017. Genomes OnLine Database (GOLD) v.6: data updates and feature enhancements. *Nucleic Acids Res* 45:D446–D456. <https://doi.org/10.1093/nar/gkw992>.
5. 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>.
6. Huntemann M, Ivanova NN, Mavromatis K, Tripp HJ, Paez-Espino D, Palaniappan K, Szeto E, Pillay M, Chen IM-a, Pati A, Nielsen T, Markowitz VM, Kyrpides NC. 2015. The standard operating procedure of the DOE-JGI microbial genome annotation pipeline (MGAP v.4). *Stand Genomic Sci* 10:86. <https://doi.org/10.1186/s40793-015-0077-y>.
7. 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.
8. Pruesse E, Quast C, Knittel FB, 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. <https://doi.org/10.1093/nar/gkm864>.
9. Markowitz VM, Mavromatis K, 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. <https://doi.org/10.1093/bioinformatics/btp393>.
10. Chen IA, Markowitz VM, Chu K, Palaniappan K, Szeto E, Pillay M, Ratner A, Huang J, Andersen E, Huntemann M, Varghese N, Hadjithomas M, Tennessen K, Nielsen T, Ivanova NN, Kyrpides NC. 2017. IMG/M: integrated genome and metagenome comparative data analysis system. *Nucleic Acids Res* 45:D507–D516. <https://doi.org/10.1093/nar/gkw929>.