



# Draft Genome Sequence of *Monaibacterium marinum* C7<sup>T</sup>, Isolated from Seawater from the Menai Straits, Wales, United Kingdom

Tatyana N. Chernikova,<sup>a</sup> Nikos Kyrpides,<sup>c</sup> Rafael Bargiela,<sup>a</sup> Tanja Woyke,<sup>c</sup> Nicole Shapiro,<sup>c</sup> William B. Whitman,<sup>b</sup> Peter N. Golyshin<sup>a</sup>

<sup>a</sup>School of the Biological Sciences, Bangor University, Bangor, Gwynedd, United Kingdom

<sup>b</sup>Department of Microbiology, University of Georgia, Athens, Georgia, USA

<sup>c</sup>DOE, Joint Genome Institute, Walnut Creek, California, USA

**ABSTRACT** Here, we report the draft genome sequence of *Monaibacterium marinum* C7<sup>T</sup>, a strain that represents a new member of the *Roseobacter* clade of the family *Rhodobacteraceae* (*Alphaproteobacteria*). The genome size of *Monaibacterium marinum* C7<sup>T</sup> is 3.7 Mb (3,734,267 bp), with a G+C content of 58.86%.

The *Roseobacter* clade within the *Rhodobacteraceae* (*Alphaproteobacteria*) represents a distinct and diverse group of microorganisms that constitute up to 25% of the coastal marine bacteria (1, 2). The study of the genomic features of members of this clade provides an insight into their ecology, physiology, niche specificity, and adaptation to the environmental stimuli.

The new strain *Monaibacterium marinum* C7<sup>T</sup> was isolated from a seawater sample taken from Menai Straits (Wales, UK) and represents a distinct lineage within the *Roseobacter* clade of the family *Rhodobacteraceae* within the *Alphaproteobacteria* (3). Strain C7<sup>T</sup> is an aerobic Gram-negative mesophilic nonmotile short-rod-shaped bacterium that can grow under microaerophilic conditions. It grows on ONR7a medium (4) in the range of NaCl concentrations of 0 to 9% (wt/vol), at temperatures between 4 and 31°C, and in pH range 5.5 to 9.0 with the optimum pH of 7.5.

The DNA was isolated using Meta-G-Nome DNA isolation kit (Epicentre Biotechnologies), according to the manufacturer's instructions.

The genome of *Monaibacterium marinum* C7<sup>T</sup> was sequenced in the framework of the project Genomic Encyclopedia of Bacteria and Archaea, phase III, at the Department of Energy (DOE) Joint Genome Institute (JGI) using Illumina technology (5). Sequencing of a standard shotgun library with 300-bp inserts using the Illumina 9 HiSeq 2000 1-Tb platform generated 6,934,098 reads totaling 1,040.1 Mbp of sequence. Known Illumina artifacts were removed, and PhiX sequencing was done by filtering all raw Illumina sequence data using BBDuk (6). Reads with fewer than 51 bp after trimming were discarded from the analysis. The remaining reads were subjected to sequence mapping and masking using BBMAP (6) and BBMask (6), respectively. Assembly was performed using SPAdes version 3.6.2 (7). Annotation was done using the JGI-IMG Annotation Pipeline version 4.14.1.

The final draft assembly contained 27 contigs in 27 scaffolds, totaling 3.735 Mbp in size, and was based on 1,040.1 Mbp of Illumina data with a mapped coverage of 279.2-fold.

The largest contig was 1,171.7 kb. The total number of genes was 3,615, including 57 RNA genes (rRNA, 6; 5S RNA, 4; 16S RNA, 1; 23S RNA, 1; tRNA, 42; other RNA genes 3). Moreover, annotation revealed 3,558 protein-coding genes, with 2,915 having functional predictions. Analysis of the draft genome sequence of *Monaibacterium*

Received 20 November 2017 Accepted 27 November 2017 Published 1 February 2018

**Citation** Chernikova TN, Kyrpides N, Bargiela R, Woyke T, Shapiro N, Whitman WB, Golyshin PN. 2018. Draft genome sequence of *Monaibacterium marinum* C7<sup>T</sup>, isolated from seawater from the Menai Straits, Wales, United Kingdom. *Genome Announc* 6:e01444-17. <https://doi.org/10.1128/genomeA.01444-17>.

**Copyright** © 2018 Chernikova 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 Tatyana N. Chernikova, [t.chernikova@bangor.ac.uk](mailto:t.chernikova@bangor.ac.uk).

*marinum* C7<sup>T</sup> using AromaDeg (8) and BLASTP against an in-house database (9), which contains experimentally characterized enzymes involved in the catabolism of aliphatic and aromatic compounds, predicted genes encoding two alkane hydroxylases and two P450 cytochromes (albeit, this strain was not able to grow on aliphatic hydrocarbons of carbon chain length 10 to 20 [3]), one catechol 1,2 dioxygenase for the *ortho*-pathway, and three carboxymuconate cycloisomerases. In addition, the gene inspection using antiSMASH (10) revealed the presence of gene clusters for the biosynthesis of terpenes, homoserine lactones, aryl propylene, and ladderanes.

**Accession number(s).** This draft genome sequence of *Monaibacterium marinum* C7<sup>T</sup> has been deposited in DDBJ/ENA/GenBank under the accession no. [OCTN01000001](https://doi.org/10.1128/AEM.71.10.5665-5677.2005) to [OCTN01000027](https://doi.org/10.1128/AEM.71.10.5665-5677.2005). The versions described in this paper are the first versions, OCTN01000001.1 to OCTN01000027.1.

## ACKNOWLEDGMENTS

This work was conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, and was supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-05CH11231, by an EU Horizon 2020 Research and Innovation Program Grant INMARE (contract 634486) (to T.N.C. and P.N.G.). R.B. acknowledges the support of the Supercomputing Wales project, which is partially funded by the European Regional Development Fund (ERDF) via the Welsh Government.

## REFERENCES

- Buchan A, González JM, Moran MA. 2005. Overview of the marine *Roseobacter* lineage. *Appl Environ Microbiol* 71:5665–5677. <https://doi.org/10.1128/AEM.71.10.5665-5677.2005>.
- Brinkhoff T, Giebel HA, Simon M. 2008. Diversity, ecology, and genomics of the *Roseobacter* clade: a short overview. *Arch Microbiol* 189:531–539. <https://doi.org/10.1007/s00203-008-0353-y>.
- Chernikova TN, Dallimore J, Lünsdorf H, Heipieper HJ, Golyshin PN. 2017. *Monaibacterium marinum*, gen. nov., sp. nov., a new member of the *Alphaproteobacteria* isolated from seawater of Menai Straits, Wales, UK. *Int J Syst Evol Microbiol* 67:3310–3317. <https://doi.org/10.1099/ijsem.0.002111>.
- Dyksterhouse SE, Gray JP, Herwig RP, Lara JC, Staley JT. 1995. *Cycloclasticus pugetii* gen. nov., sp. nov., an aromatic hydrocarbon-degrading bacterium from marine sediments. *Int J Syst Bacteriol* 45:116–123. <https://doi.org/10.1099/00207713-45-1-116>.
- Bennett S. 2004. Solexa Ltd. *Pharmacogenomics* 5:433–438. <https://doi.org/10.1517/14622416.5.4.433>.
- Bushnell S, Budde J, Catino T, Cole J, Derti A, Kelso R, Collins ML, Molino G, Sheridan P, Monahan J, Urdea M. 1999. ProbeDesigner: for the design of probe sets for branched DNA (bDNA) signal amplification assays. *Bioinformatics* 15:348–355. BBTools Software Package, URL <http://sourceforge.net/projects/bbmap>.
- 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>.
- Duarte M, Jauregui R, Vilchez-Vargas R, Junca H, Pieper DH. 2014. AromaDeg, a novel database for phylogenomics of aerobic bacterial degradation of aromatics. *Database* 2014:bau118. <https://doi.org/10.1093/database/bau118>.
- Guazzaroni ME, Herbst FA, Lores I, Tamames J, Peláez AI, López-Cortés N, Alcaide M, Del Pozo MV, Vieites JM, von Bergen M, Gallego JL, Bargiela R, López-López A, Pieper DH, Rosselló-Móra R, Sánchez J, Seifert J, Ferrer M. 2013. Metaproteogenomic insights beyond bacterial response to naphthalene exposure and bio-stimulation. *ISME J* 7:122–136. <https://doi.org/10.1038/ismej.2012.82>.
- Blin K, Medema MH, Kottmann R, Lee SY, Weber T. 2017. The antiSMASH database, a comprehensive database of microbial secondary metabolite biosynthetic gene clusters. *Nucleic Acids Res* 45:D555–D559. <https://doi.org/10.1093/nar/gkw960>.