AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

# Metagenome-Assembled Genome Sequence of Marine *Rhizobiaceae* sp. Strain MnEN-MB40S, Obtained from Manganese-Oxidizing Enrichment Culture

Masataka Aoki,[a,b] Nozomi Nakahara,[c] Masataka Kusube,[d] Kazuaki Syutsubo[a,e]

[a]Regional Environment Conservation Division, National Institute for Environmental Studies, Tsukuba, Ibaraki, Japan
[b]Department of Civil Engineering, National Institute of Technology, Wakayama College, Gobo, Wakayama, Japan
[c]Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki, Japan
[d]Department of Applied Chemistry and Biochemistry, National Institute of Technology, Wakayama College, Gobo, Wakayama, Japan
[e]Research Center for Water Environment Technology, School of Engineering, the University of Tokyo, Bunkyo, Tokyo, Japan

**ABSTRACT** Here, we report a new metagenome-assembled genome (MAG) from a marine *Rhizobiaceae* species. The MnEN-MB40S genome was assembled from a manganese-oxidizing enrichment culture metagenome. A 4.1-Mb MAG comprising 26 contigs, with a GC content of 60.0%, was obtained. This MAG contributes to the genomic information regarding the family *Rhizobiaceae*.

The family *Rhizobiaceae* comprises genera that are mainly associated with soil and plant hosts (1). Some *Rhizobiaceae* genera can potentially be used for bioremediation of heavy metals and biodegradation of toxic compounds (1). However, their metabolic capabilities and ecological roles, particularly in marine environments, remain unclear because of the limited availability of cultured marine isolates and their sequenced genomes.

Here, we report a new *Rhizobiaceae*-associated metagenome-assembled genome (MAG) that was derived from a marine manganese-oxidizing enrichment culture originating from seawater (water depth, 0 m) in the Nada coastal area of Wakayama, Japan (33°49′52.6″N, 135°10′31.8″E) (2). Briefly, seawater (100 mL), used as an enrichment inoculum, was collected directly in a 500-mL glass bottle. A 1-L-scale enrichment experiment was performed using polycaprolactone (as a solid organic substrate) and a MARINE ART SF-1 artificial seawater (Osaka Yakken Co., Ltd.)-based medium under aerobic conditions at ~25°C (see reference 2 for details). After 3,340 h of enrichment, 22.5 mL of the planktonic fraction of the enrichment culture was collected using a 25-mL pipette and stored as a glycerol stock at −80°C until DNA extraction. Total genomic DNA for metagenomic sequencing was extracted using an MPure bacterial DNA extraction kit (MP Biomedicals). A sequencing library was prepared using an MGIEasy FS DNA library prep set (MGI Tech Co., Ltd.) and an MGIEasy DNA adapters-96 (plate) kit (MGI). Circularized DNA and DNA nanoballs were prepared using an MGIEasy circularization kit (MGI) and a DNBSEQ-G400RS high-throughput sequencing kit (MGI), respectively. Paired-end (2 × 200-bp) sequencing on a DNBSEQ-G400 sequencer (MGI) produced 135,388,150 raw sequence reads (total of 27,077,630,000 bp). Default parameters were used for all software unless otherwise specified. Adapter sequences were removed, and the raw sequence reads were quality filtered using fastp v0.20.0 (quality scores of ≥30) (3). Quality-filtered reads were assembled using MEGAHIT v1.2.9 (4). Assembled contigs were binned using MaxBin 2.0 v2.2.7 (5), MetaBAT 2 v2.12.1 (6), and MyCC vMyCC_2017 (7) and then refined using DAS Tool v1.1.3 (8). Genes were predicted using the DFAST pipeline v1.2.16 (https://dfast.ddbj.nig.ac.jp) (9). Genome coverage was calculated using Minimap2 v2.23 (10). The EzBioCloud 16S rRNA gene-based identification service (database v2021.07.07) (11) was used for the 16S rRNA gene similarity search. Taxonomic affiliation was determined using GTDB-Tk v1.7.0 and the Genome Taxonomy Database (GTDB) R202 (12) in KBase (13). MAG

**Editor** J. Cameron Thrash, University of Southern California



Address correspondence to Masataka Aoki, aoki.masataka@nies.go.jp.

The authors declare no conflict of interest.

**Received** 6 July 2022
**Accepted** 1 August 2022
**Published** 17 August 2022

**TABLE 1** Genome statistics and taxonomic information for *Rhizobiaceae* MAG MnEN-MB40S

| Parameter | Finding |
|---|---|
| Genome size (bp) | 4,129,535 |
| No. of contigs | 26 |
| GC content (%) | 60.0 |
| $N_{50}$ (bp) | 294,856 |
| Genome coverage (×) | 59 |
| No. of coding sequences | 3,879 |
| No. of rRNA genes (5S, 16S, 23S) | 1 (0, 1, 0) |
| No. of tRNA genes | 32 |
| Completeness (%) | 75.34 |
| Contamination (%) | 0.00 |
| Taxonomic affiliation | |
| Phylum level | *Proteobacteria* |
| Class level | *Alphaproteobacteria* |
| Order level | *Rhizobiales* |
| Family level | *Rhizobiaceae* |
| Genus level | SPNT01 |
| Closest cultivated relative based on 16S rRNA gene sequence similarity | *Hoeflea prorocentri* PM5-8 |
| Similarity to closet cultivated relative (%) | 97.58 |
| GenBank accession no. for closet cultivated relative | KY264918 |

completeness and contamination were calculated using CheckM (marker set, *Alphaproteobacteria*) (14) in DFAST.

Table 1 summarizes the genome statistics and taxonomic information of *Rhizobiaceae* MAG MnEN-MB40S. The sequence similarity of the 16S rRNA gene to the closest cultivated relative (97.58%) was below the proposed species boundary (15). A BLASTp search (E value threshold, $1 \times 10^{-20}$; query coverage, ≥70%; sequence identity, ≥40%) (16) against nonredundant sequences in NCBI and UniProtKB/Swiss-Prot revealed the presence of putative multicopper oxidase (MCO) genes [related to Mn(II) oxidation]. The locus tags MnENMB40S_08170 and MnENMB40S_10010 were homologous to the MCO gene *moxA* from *Pedomicrobium* sp. strain ACM 3067 (GenBank accession number CAJ19378) (71.8% sequence identity) (17) and the MCO gene *cueO* from *Escherichia coli* K-12 MG1655 (GenBank accession number P36649) (44.6% sequence identity) (18), respectively.

**Data availability.** The annotated MAG of *Rhizobiaceae* sp. strain MnEN-MB40S is available in DDBJ/EMBL-Bank/GenBank under the accession numbers BRLC01000001 to BRLC01000026. The raw DNBSeq read data have been deposited in the Sequence Read Archive (SRA) under the accession number DRX364682.

## ACKNOWLEDGMENTS

We thank Yukina Miyashita for her assistance in this study.

This work was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI (grant 20K15222). The computations were partially performed using the NIG supercomputer at the ROIS National Institute of Genetics.

We declared that there are no competing interests.

## REFERENCES

1. Carareto Alves LM, de Souza JAM, Varani AdM, Lemos EGdM. 2014. The family *Rhizobiaceae*, p 419–437. *In* Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F (ed), The prokaryotes, Springer, Berlin, Germany. https://doi.org/10.1007/978-3-642-30197-1_297.
2. Aoki M, Miyashita Y, Tran PT, Okuno Y, Watari T, Yamaguchi T. 2021. Enrichment of marine manganese-oxidizing microorganisms using polycaprolactone as a solid organic substrate. Biotechnol Lett 43:813–823. https://doi.org/10.1007/s10529-021-03088-z.
3. Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890. https://doi.org/10.1093/bioinformatics/bty560.
4. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct *de Bruijn* graph. Bioinformatics 31:1674–1676. https://doi.org/10.1093/bioinformatics/btv033.
5. Wu Y-W, Simmons BA, Singer SW. 2016. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics 32:605–607. https://doi.org/10.1093/bioinformatics/btv638.
6. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z. 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ 7:e7359. https://doi.org/10.7717/peerj.7359.

7. Lin H-H, Liao Y-C. 2016. Accurate binning of metagenomic contigs via automated clustering sequences using information of genomic signatures and marker genes. Sci Rep 6:24175. https://doi.org/10.1038/srep24175.
8. Sieber CMK, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG, Banfield JF. 2018. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. Nat Microbiol 3:836–843. https://doi.org/10.1038/s41564-018-0171-1.
9. Tanizawa Y, Fujisawa T, Nakamura Y. 2018. DFAST: a flexible prokaryotic genome annotation pipeline for faster genome publication. Bioinformatics 34:1037–1039. https://doi.org/10.1093/bioinformatics/btx713.
10. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34:3094–3100. https://doi.org/10.1093/bioinformatics/bty191.
11. Yoon S-H, Ha S-M, Kwon S, Lim J, Kim Y, Seo H, Chun J. 2017. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. Int J Syst Evol Microbiol 67:1613–1617. https://doi.org/10.1099/ijsem.0.001755.
12. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2020. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics 36:1925–1927. https://doi.org/10.1093/bioinformatics/btz848.
13. Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, Dehal P, Ware D, Perez F, Canon S, Sneddon MW, Henderson ML, Riehl WJ, Murphy-Olson D, Chan SY, Kamimura RT, Kumari S, Drake MM, Brettin TS, Glass EM, Chivian D, Gunter D, Weston DJ, Allen BH, Baumohl J, Best AA, Bowen B, Brenner SE, Bun CC, Chandonia J-M, Chia J-M, Colasanti R, Conrad N, Davis JJ, Davison BH, DeJongh M, Devoid S, Dietrich E, Dubchak I, Edirisinghe JN, Fang G, Faria JP, Frybarger PM, Gerlach W, Gerstein M, Greiner A, Gurtowski J, Haun HL, He F, Jain R, et al. 2018. KBase: the United States Department of Energy Systems Biology Knowledgebase. Nat Biotechnol 36:566–569. https://doi.org/10.1038/nbt.4163.
14. 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.
15. Kim M, Oh H-S, Park S-C, Chun J. 2014. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int J Syst Evol Microbiol 64:346–351. https://doi.org/10.1099/ijs.0.059774-0.
16. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402. https://doi.org/10.1093/nar/25.17.3389.
17. Ridge JP, Lin M, Larsen EI, Fegan M, McEwan AG, Sly LI. 2007. A multicopper oxidase is essential for manganese oxidation and laccase-like activity in *Pedomicrobium* sp. ACM 3067. Environ Microbiol 9:944–953. https://doi.org/10.1111/j.1462-2920.2006.01216.x.
18. Su J, Deng L, Huang L, Guo S, Liu F, He J. 2014. Catalytic oxidation of manganese(II) by multicopper oxidase CueO and characterization of the biogenic Mn oxide. Water Res 56:304–313. https://doi.org/10.1016/j.watres.2014.03.013.