





# Draft Genome Sequence of Novel Filterable *Rhodospirillales* Bacterium Strain TMPK1, Isolated from Soil

 Ryosuke Nakai,<sup>a</sup> Hiroyuki Kusada,<sup>b</sup> Fumihiko Sassa,<sup>c</sup> Susumu Morigasaki,<sup>d</sup> Hisayoshi Hayashi,<sup>d</sup> Naoki Takaya,<sup>d,e</sup>  
 Hideyuki Tamaki<sup>b,d,e</sup>

<sup>a</sup>Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Sapporo, Hokkaido, Japan

<sup>b</sup>Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Ibaraki, Japan

<sup>c</sup>Department of Electronics, Graduate School of Information Science and Electrical Engineering, Kyushu University, Fukuoka, Fukuoka, Japan

<sup>d</sup>Faculty of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki, Japan

<sup>e</sup>Microbiology Research Center for Sustainability, University of Tsukuba, Tsukuba, Ibaraki, Japan

Ryosuke Nakai and Hiroyuki Kusada contributed equally to this work. Author order was determined by their contribution to the experiments and manuscript preparation.

**ABSTRACT** We report the draft genome sequence of novel *Rhodospirillales* bacterium strain TMPK1, isolated from a micropore-filtered soil suspension. This strain has a genome of 4,249,070 bp, comprising 4,151 protein-coding sequences. The genome sequence data further suggest that strain TMPK1 is an alphaproteobacterium capable of carotenoid production.

The order *Rhodospirillales* (class *Alphaproteobacteria*) is a metabolically diverse group, including acetic acid, photosynthetic purple, and magnetotactic bacteria (1, 2). Here, we report the draft genome sequence of novel *Rhodospirillales* bacterium strain TMPK1.

Strain TMPK1 was isolated from upland soil at the Tsukuba-Plant Innovation Research Center (T-PIRC), University of Tsukuba, Ibaraki, Japan. Briefly, ~3 g soil was suspended in 27 ml UltraPure DNase/RNase-free distilled water (Thermo Fisher, Tokyo, Japan), and the suspension was filtered through a sterile 0.45- $\mu$ m-pore-size filter (Millex-HV syringe filter unit; Merck Millipore, Tokyo, Japan) to target novel filterable bacteria (reviewed in reference 3). The filtered solution was spread on a gel-filled microwell array device (76 by 26 mm; pitch, 1.0 mm; well depth, 0.8 mm) composed of 900 miniature chambers for culturing microbial cells (F. Sassa, T. Kiyokawa, A. S. Utada, K. Nagata, R. Mogi, M. Hamada, T. Inaba, N. Obana, M. Yokokawa, M. Toyofuku, H. Suzuki, and N. Nomura, unpublished data) filled with 1/100-strength tryptic soy agar medium. The microcolonies formed on the device were transferred onto agar plates containing the same ingredients. Strain TMPK1 was purified by single-colony isolation.

For genome sequencing, strain TMPK1 was cultured in the R2A broth “DAIGO” (Nihon Pharmaceutical Co., Ltd., Tokyo, Japan) at 25°C for 2 weeks. DNA extraction was performed using Genomic-tip 100/G columns (Qiagen, Tokyo, Japan). A sequence library was constructed using the NEBNext Ultra II DNA library prep kit for Illumina (New England BioLabs, Tokyo, Japan) and then sequenced via 150-bp paired-end sequencing using a NovaSeq device (Illumina, Tokyo, Japan). For subsequent bioinformatic analysis, the default settings were used for all software unless otherwise specified. Raw reads (total, 10,997,950 reads;  $\approx$  1.6 Gbp) were trimmed using fastp v. 0.20.0 (4) with the settings -q 5, -u 50, -n 15, and -l 150 and discarded by the in-house pipeline of Novogene Co., Ltd. (Beijing, China), according to the following criteria: reads containing (i) adapter sequences, (ii) > 10% uncertain nucleotide bases, or (iii) over 50% low-quality bases (Q score  $\leq$  5). The remaining sequence reads were assembled using SPAdes v. 3.13.0 (5) (--careful, -k 21,33,55,77,99,127, --cov-cutoff auto). The assembled contigs were processed using the DFAST annotation pipeline v. 1.2.11 (6). The genome map with Clusters of Orthologous Groups (COGs) was analyzed and visualized using the whole-genome analysis pipeline of EzBioCloud (ChunLab, Inc., Seoul, South Korea

**Citation** Nakai R, Kusada H, Sassa F, Morigasaki S, Hayashi H, Takaya N, Tamaki H. 2021. Draft genome sequence of novel filterable *Rhodospirillales* bacterium strain TMPK1, isolated from soil. *Microbiol Resour Announc* 10:e00393-21. <https://doi.org/10.1128/MRA.00393-21>.

**Editor** David A. Baltus, University of Arizona

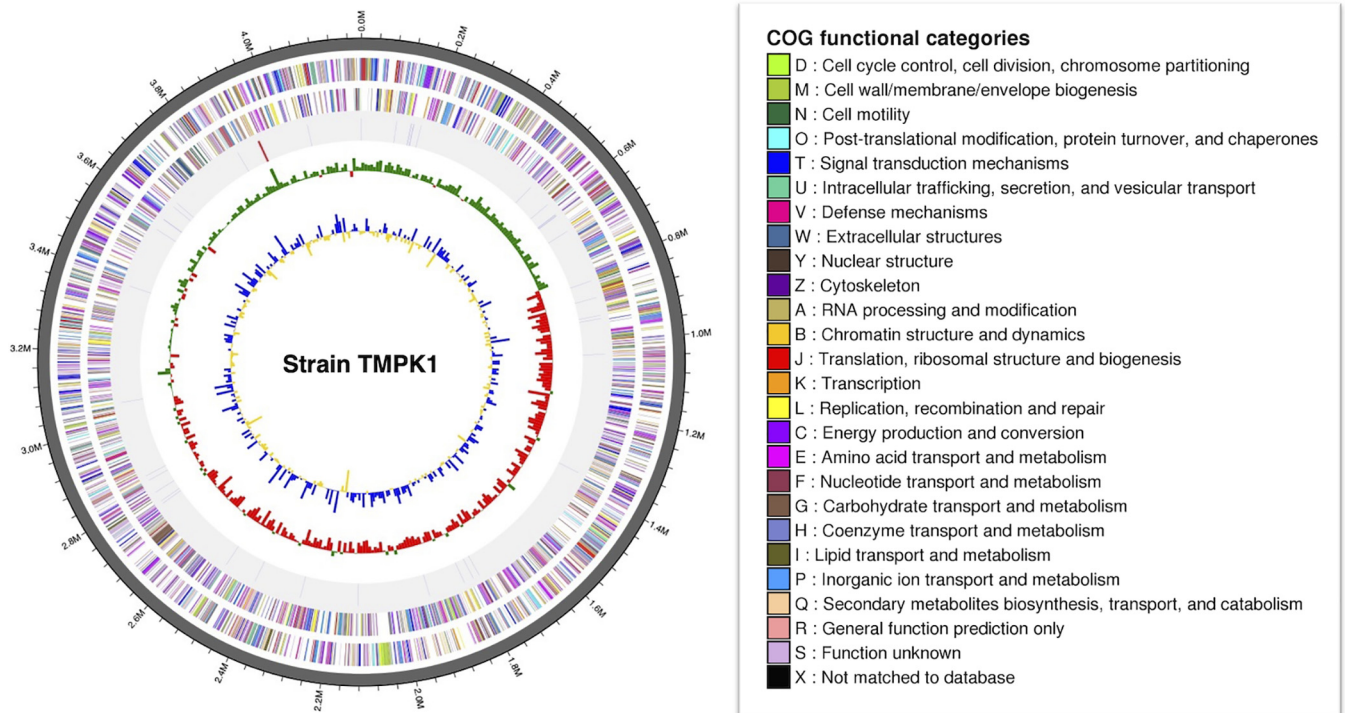
**Copyright** © 2021 Nakai et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Ryosuke Nakai, nakai-ryosuke@aist.go.jp, or Hideyuki Tamaki, tamaki-hideyuki@aist.go.jp.

**Received** 19 April 2021

**Accepted** 11 June 2021

**Published** 15 July 2021



**FIG 1** Genome map of *Rhodospirillales* bacterium strain TMPK1. The map of the longest contig (4,229,965 bp) is shown. From the circumference to the center are the annotated genes in the forward and reverse strands (the colors indicate COG functional categories [right panel]), rRNAs and tRNAs, guanine-cytosine (GC) skew metric (the mean GC-skew value is used as a baseline, with the values higher and lower than the average shown in green and red, respectively), and GC ratio metric (the mean GC ratio is used as a baseline, with the values higher and lower than the average shown in blue and yellow, respectively).

[7]). The full-length 16S rRNA gene sequence obtained was BLASTn searched against the NCBI nucleotide/nonredundant (nt/nr) database (accessed 1 June 2021). The average nucleotide identity (ANI) of strain TMPK1 against the closest type strains identified using BLASTn was calculated using the ANI calculator (8).

The TMPK1 genome comprises 3 contigs, 4,249,070 bp in total, 63.7% G+C content, 4,151 protein-coding sequences, 3 rRNA genes, and 49 tRNA genes, but no CRISPR. The draft genome was recovered at 381.06 $\times$ . The  $N_{50}$  length was 4,229,965 bp. A detailed genome map is shown in Fig. 1. Strain TMPK1 is affiliated with the order *Rhodospirillales* but shares low 16S rRNA gene sequence identities (<91%) and low ANI values (<70%), with the closest type strains being *Skermanella pratensis* W17<sup>T</sup> (9) and *Haematospirillum jordaniae* H5569<sup>T</sup> (10) within the family *Rhodospirillaceae*. The other closest taxonomically undescribed isolate and uncultivated bacterium were identified as *Rhodospirillales* bacterium SC-11 (GenBank accession no. [LC602157](#); 16S rRNA gene sequence identity, 99.3%) and a potassium mine soil clone (GenBank accession no. [JF833841](#); 94.0%), respectively. These results suggest that strain TMPK1 belongs to the phylogenetically novel *Alphaproteobacteria*. Moreover, the TMPK1 genome harbors genes for carotenoid biosynthesis (e.g., phytoene desaturase gene), similar to the genomes of several *Rhodospirillales* members. The genome data presented here serve for further elucidating the phylogenetic placement and eco-physiological potentials of this strain.

**Data availability.** The TMPK1 genome sequence was deposited in the DDBJ/ENA/GenBank database under accession no. [BOPV01000001.1](#) to [BOPV01000003.1](#) (BioProject/BioSample no. [PRJDB11251/SAMD00281908](#), DDBJ Sequence Read Archive [DRA] study/run no. [DRP007157/DRR286848](#)).

#### ACKNOWLEDGMENTS

We thank the technical staff, Kazuto Yokoyama, Kiyoshi Karube, Hironori Naoi, and Akira Saito, who have managed and operated the upland field at the Smart and Sustainable Agriculture Research Division (T-PIRC Farm), Tsukuba-Plant Innovation Research Center (T-PIRC) of the University of Tsukuba.

This study was funded by a MEXT KAKENHI Grant-in-Aid for Scientific Research on Innovative Areas “Post-Koch Ecology” (JP19H05683 to R.N., JP19H05680 to F.S., JP19H05679 to N.T., and JP19H05687 to N.T.). This study was also partly supported by JST ERATO grant no. JPMJER1502.

## REFERENCES

1. Baldani JI, Videira SS, dos Santos Teixeira KR, Reis VM, de Oliveira ALM, Schwab S, de Souza EM, Pedraza RO, Baldani VLD, Hartmann A. 2014. The family *Rhodospirillaceae*, p 533–618. In Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F (ed), *The prokaryotes: Alphaproteobacteria and Betaproteobacteria*. Springer, Berlin, Germany.
2. Hördt A, López MG, Meier-Kolthoff JP, Schleuning M, Weinhold L-M, Tindall BJ, Gronow S, Kyrpides NC, Woyke T, Göker M. 2020. Analysis of 1,000+ type-strain genomes substantially improves taxonomic classification of *Alphaproteobacteria*. *Front Microbiol* 11:468. <https://doi.org/10.3389/fmicb.2020.00468>.
3. Nakai R. 2020. Size matters: ultra-small and filterable microorganisms in the environment. *Microbes Environ* 35:ME20025. <https://doi.org/10.1264/jsme2.ME20025>.
4. Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ pre-processor. *Bioinformatics* 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>.
5. 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>.
6. 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>.
7. 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>.
8. Yoon S-H, Ha S, Lim J, Kwon S, Chun J. 2017. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* 110:1281–1286. <https://doi.org/10.1007/s10482-017-0844-4>.
9. Guo Q, Zhou Z, Zhang L, Zhang C, Chen M, Wang B, Lin M, Wang W, Zhang W, Li X. 2020. *Skermanella pratensis* sp. nov., isolated from meadow soil, and emended description of the genus *Skermanella*. *Int J Syst Evol Microbiol* 70:1605–1609. <https://doi.org/10.1099/ijsem.0.003944>.
10. Humrighouse BW, Emery BD, Kelly AJ, Metcalfe MG, Mbizo J, McQuiston JR. 2016. *Haematospirillum jordaniae* gen. nov., sp. nov., isolated from human blood samples. *Antonie Van Leeuwenhoek* 109:493–500. <https://doi.org/10.1007/s10482-016-0654-0>.