

Draft genome sequence of *Flavobacterium* sp. MAHUQ-50 isolated from the rhizosphere of banana in Magura, Bangladesh

Md. Amdadul Huq,¹ Shahina Akter,² M. Mizanur Rahman,³ Md. Anowar Khasru Parvez,⁴ Md. Morshedul Alam,⁵ Md. Shahedur Rahman⁶

AUTHOR AFFILIATIONS See affiliation list on p. 3.

ABSTRACT This study reports the draft genome sequence of *Flavobacterium* sp. MAHUQ-50, a strain isolated from the rhizosphere of banana in Magura, Bangladesh, in 2019. The genome consists of a 3,493,675 base pair chromosome assembled into 26 contigs, encoding 3,198 predicted protein coding genes.

KEYWORDS genome sequence, *Flavobacterium* sp. MAHUQ-50, rhizosphere

The genus *Flavobacterium* was established by Bergey et al. (1) and emended by Bernardet et al. (2), Kuo et al. (3), Kang et al. (4), and Dong et al. (5). Members of the genus *Flavobacterium* were isolated from different habitats, such as plant, animal, algae, seawater, freshwater, soil, rhizosphere, compost, sediment, lagoon, Antarctic lakes, glacier, activated sludge, water reservoir, and wastewater treatment plant sources (6–8). During investigations on the biodiversity of bacteria in the rhizosphere of banana, a new strain of genus *Flavobacterium*, designated as *Flavobacterium* sp. MAHUQ-50 was isolated. The soil attaching to the plant roots, termed as rhizosphere, is a living place for diverse soil microorganisms. Studying bacterial biodiversity in the plant rhizosphere is crucial because these bacteria play vital roles in plant health, nutrient acquisition, and overall ecosystem functioning (9). Genome-based data have become an essential resource in taxonomic and functional studies of microorganisms. Sequencing bacterial genomes serve as a valuable technique with broad applications in medicine, public health, and evolutionary research. It enables comprehensive analysis of bacteria, identification of biosynthetic gene cluster, detection of virulence-associated genes, monitoring of disease spread, and supports the creation of novel therapeutic agents and vaccines (10–13). This study presents the draft genome assembly of *Flavobacterium* sp. MAHUQ-50.

Strain MAHUQ-50 was isolated from the rhizosphere of banana, located in Dighalgram, Magura, Bangladesh (GPS map: 23.410994, 89.334212). One gram of soil sample was suspended in 9 mL of sterile 0.85% (w/v) NaCl solution. The suspension was serially diluted up to a 10⁻⁶ dilution, and 100 µL of individual dilution was spread onto R2A agar plates. Then, the plates were placed in a 30°C incubator for 3 days. Single colonies were purified by repeated streaking on fresh R2A agar plates. The genomic DNA was extracted from the bacterial culture that was grown in R2A broth for 24 h using a genomic DNA extraction kit according to the manufacturer's protocol (Solgent, Republic of Korea). Whole-genome sequencing was performed at the World Data Center for Microorganisms (WDCM), The Institute of Microbiology of the Chinese Academy of Sciences, Beijing 100101, China. The draft genome sequence of the strain MAHUQ-50 was determined using an Illumina HiSeq 3000 platform. The DNA fragments were size selected using AMPure XP beads to isolate those within the desired range, typically 200–300 bp for standard Illumina sequencing libraries. The library for sequencing was prepared using the Nextera XT DNA Library Prep Kit (Illumina, San Diego, USA). After

Editor David A. Baltrus, The University of Arizona, Tucson, Arizona, USA

Address correspondence to Md. Shahedur Rahman, ms.rahman@just.edu.bd.

The authors declare no conflict of interest.

Received 9 June 2025

Accepted 1 July 2025

Published 31 July 2025

Copyright © 2025 Huq 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/).

TABLE 1 Genomic features of *Flavobacterium* sp. MAHUQ-50

Description of source	
Location	Magura, Bangladesh
Time	2019
Type	Rhizosphere of banana
Sequencing summary	
Coverage	248×
Total bases	3,493,675
GC content	35.5%
Assembly report	
Contigs	26
Contig L_{50}	3
Contig N_{50}	468.1 kb
Genome length	3.5 Mb
Annotation report	
Genes (total)	3,258
CDSs (total)	3,205
CDSs (with protein)	3,198
ncRNAs	3
tRNA	46
Quality analysis	
Completeness	99.22%
Contamination	0.16%

library preparation, sequencing was carried out using the Illumina HiSeq 3000 platform with paired-end 2×150 bp reads following the standard protocol provided by the manufacturer. Illumina's bcl2fastq software version 2.20.0 was utilized with default parameters to demultiplex the raw reads and convert them into FASTQ format. For further downstream analysis, raw paired-end reads from the HiSeq platform were used. Reads were processed with Trimmomatic version 0.38 (14). The genome sequence was assembled by SOAPdenovo v. v2.04; SPAdes v. v3.13.0 assembler (15). The genome was annotated with the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) version 6.3 (16–18). The strain was identified using a whole genome-based taxonomic analysis in the Type (Strain) Genome Server (<https://tygs.dsmz.de>). Default parameters were used for all software unless otherwise specified. The genome-to-genome distance (GGDH) bioinformatics tool (19) was used to measure *in silico* DNA-DNA hybridization (isDDH) values. The best matches to the strain MAHUQ-50 genome are the type strains *Flavobacterium amniphilum* KYPY10 (accession [NZ_JAMLJL000000000](#)) and *Flavobacterium terrae* DSM 18829 (accession [GCF_900142035](#)), with *in silico* DNA-DNA hybridization values of 22.3% and 21.0%, respectively.

The genome of *Flavobacterium* sp. MAHUQ-50 is 3,493,675 bp long with a GC content of 35.5%. The assembly contains 26 contigs with a coverage of 248×. A total of 3,258 genes were predicted, of which 3,198 are protein-coding genes. The details of genomic features are presented in Table 1.

ACKNOWLEDGMENTS

No external fund was received for this study.

Md. Amdadul Huq, Conceptualization, Formal analysis, Writing – original draft, Writing – review and editing | Shahina Akter, M. Mizanur Rahman, Md. Anowar Khasru Parvez, and Md. Morshedul Alam, Formal analysis, | Md. Shahedur Rahman, Writing – review and editing.

AUTHOR AFFILIATIONS

¹Department of Life Sciences, College of BioNano Technology, Gachon University, Seongnam, Republic of Korea

²Department of Food Science and Biotechnology, Gachon University, Seongnam, Republic of Korea

³Department of Biotechnology and Genetic Engineering, Faculty of Biological Science, Islamic University, Kushtia, Bangladesh

⁴Department of Microbiology, Jahangirnagar University, Dhaka, Bangladesh

⁵Department of Biochemistry and Microbiology, School of Health and Life Sciences, North South University, Dhaka, Bangladesh

⁶Department of Genetic Engineering and Biotechnology, Jashore University of Science and Technology, Jashore, Bangladesh

AUTHOR ORCIDS

Md. Amdadul Huq  <http://orcid.org/0000-0002-4831-6368>

Md. Shahedur Rahman  <http://orcid.org/0009-0005-2028-4885>

AUTHOR CONTRIBUTIONS

Md. Amdadul Huq, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Writing – original draft, Writing – review and editing | Shahina Akter, Formal analysis | M. Mizanur Rahman, Formal analysis | Md. Anowar Khasru Parvez, Formal analysis | Md. Morshedul Alam, Formal analysis | Md. Shahedur Rahman, Investigation, Writing – review and editing

DATA AVAILABILITY

The draft genome sequence of *Flavobacterium* sp. MAHUQ-50 was deposited in GenBank under accession number [JBOEFP000000000](#). The raw reads are available in SRA under accession number [SRR33604969](#). The version described in this paper is the first version, [JBOEFP000000000.1](#). The data have been deposited in GenBank under BioSample number [SAMN48698540](#) and BioProject number [PRJNA1267049](#). The genome assembly is available under the accession [ASM5063252v1](#).

REFERENCES

- Bergey DH, Harrison FC, Breed RS, Hammer BW, Huntoon FM. 1923. Genus II. *Flavobacterium* gen. nov., p 97–117. In Bergey's Manual of Determinative Bacteriology, 1st ed. Williams & Wilkins, Baltimore.
- Bernardet JF, Segers P, Vancanneyt M, Berthe F, Kersters K, Vandamme P. 1996. Cutting a gordian knot: emended classification and description of the genus *Flavobacterium*, emended description of the family *Flavobacteriaceae*, and proposal of *Flavobacterium* *hydatis* nom. nov. (Basonym, *Cytophaga aquatilis* Strohl and Tait 1978). *Int J Syst Bacteriol* 46:128–148. <https://doi.org/10.1099/00207713-46-1-128>
- Kuo I, Saw J, Kapan DD, Christensen S, Kaneshiro KY, Donachie SP. 2013. *Flavobacterium akiainvivens* sp. nov., from decaying wood of *Wikstroemia oahuensis*, Hawaii, and emended description of the genus *Flavobacterium*. *Int J Syst Evol Microbiol* 63:3280–3286. <https://doi.org/10.1099/ijs.0.047217-0>
- Kang JY, Chun J, Jahng KY. 2013. *Flavobacterium aciduliphilum* sp. nov., isolated from freshwater, and emended description of the genus *Flavobacterium*. *Int J Syst Evol Microbiol* 63:1633–1638. <https://doi.org/10.1099/ijs.0.044495-0>
- Dong K, Chen F, Du Y, Wang G. 2013. *Flavobacterium enshiense* sp. nov., isolated from soil, and emended descriptions of the genus *Flavobacterium* and *Flavobacterium caucense*, *Flavobacterium saliperosum* and *Flavobacterium suncheonense*. *Int J Syst Evol Microbiol* 63:886–892. <https://doi.org/10.1099/ijs.0.039974-0>
- Akter S, Lee SY, Moon B, Huq MA. 2021. *Flavobacterium agri* sp. nov., a novel bacterial species isolated from rhizospheric soil of *Coriandrum sativum*. *Arch Microbiol* 203:701–706. <https://doi.org/10.1007/s00203-020-02068-7>
- Liu Q, Siddiqi MZ, Liu Q, Huq MA, Lee SY, Choi KD, Im WT. 2018. *Flavobacterium hankyongi* sp. nov., isolated from activated sludge. *Int J Syst Evol Microbiol* 68:1732–1736. <https://doi.org/10.1099/ijsem.0.002739>
- Huq MA, Akter S, Lee SY. 2018. *Flavobacterium chungangensis* sp. nov., a bacterium isolated from soil of Chinese cabbage garden. *Curr Microbiol* 75:842–848. <https://doi.org/10.1007/s00284-018-1456-6>
- Ling N, Wang T, Kuzyakov Y. 2022. Rhizosphere bacteriome structure and functions. *Nat Commun* 13:836. <https://doi.org/10.1038/s41467-022-28448-9>
- Donkor ES. 2013. Sequencing of bacterial genomes: principles and insights into pathogenesis and development of antibiotics. *Genes (Basel)* 4:556–572. <https://doi.org/10.3390/genes4040556>
- Mustafa AS. 2024. Whole genome sequencing: applications in clinical bacteriology. *Med Princ Pract* 33:185–197. <https://doi.org/10.1159/000538002>
- Huq MA, Kim YJ, Rahman MM, Alam MM, Srinivasan S, Kang KK, Akter S. 2025. *Vogesella oryzae* sp. nov., isolated from the rhizosphere of rice and *in silico* genome mining for the prediction of biosynthetic gene clusters. *Int J Syst Evol Microbiol* 75:006687. <https://doi.org/10.1099/ijsem.0.006687>
- Huq MA, Rahman MS, Rahman MM. 2024. *Aquicola agrisoli* sp. nov., isolated from rhizospheric soil of eggplant and *in silico* genome mining

- for the prediction of biosynthetic gene clusters. *Int J Syst Evol Microbiol* 74:006355. <https://doi.org/10.1099/ijsem.0.006355>
14. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
 15. 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>
 16. Li W, O'Neill KR, Haft DH, DiCuccio M, Chetvernin V, Badretdin A, Coulouris G, Chitsaz F, Derbyshire MK, Durkin AS, Gonzales NR, Gwadz M, Lanczycki CJ, Song JS, Thanki N, Wang J, Yamashita RA, Yang M, Zheng C, Marchler-Bauer A, Thibaud-Nissen F. 2021. RefSeq: expanding the prokaryotic genome annotation pipeline reach with protein family model curation. *Nucleic Acids Res* 49:D1020–D1028. <https://doi.org/10.1093/nar/gkaa1105>
 17. Haft DH, DiCuccio M, Badretdin A, Brover V, Chetvernin V, O'Neill K, Li W, Chitsaz F, Derbyshire MK, Gonzales NR, Gwadz M, Lu F, Marchler GH, Song JS, Thanki N, Yamashita RA, Zheng C, Thibaud-Nissen F, Geer LY, Marchler-Bauer A, Pruitt KD. 2018. RefSeq: an update on prokaryotic genome annotation and curation. *Nucleic Acids Res* 46:D851–D860. <https://doi.org/10.1093/nar/gkx1068>
 18. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>
 19. Meier-Kolthoff JP, Carbasse JS, Peinado-Olarte RL, Göker M. 2022. TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acids Res* 50:D801–D807. <https://doi.org/10.1093/nar/gkab902>