AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

| Bacteriology | Announcement

Draft genome sequence of *Nocardioides* strain MAHUQ-72 isolated from soil sample of a rice field located in Anseong, South Korea

Md. Amdadul Huq,[1] Shahina Akter,[2] Md. Shahidul Islam,[3] Md. Shahedur Rahman[4]

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT This study reports the draft genome sequence of *Nocardioides* strain MAHUQ-72, a bacterial strain isolated from the soil sample of a rice field located in Anseong, South Korea, in 2020. The genome consists of 4,732,459 base pairs assembled into 23 contigs, encoding 4,555 predicted protein-coding genes.

KEYWORDS genome sequence, *Nocardioides*strain MAHUQ-72, rice field

The genus *Nocardioides* of the family *Nocardioidaceae* belongs to the phylum *Actinomycetota* and was first described by Prauser (1), currently comprises 171 validly published species (2) that have been isolated from different environments, including soil, water, plant roots, oil shale columns, meadows, deserts, forests, dandelions, sewage sludge, and caves (3). During investigations on bacterial biodiversity in a rice field, a new strain of the genus *Nocardioides,* designated *Nocardioides* strain MAHUQ-72, was isolated, and its draft genome assembly is presented in this study.

Strain MAHUQ-72 was isolated from the soil sample of a rice field (Date of collection: 2020.10.20), located in Anseong, South Korea (37° 00′ 31″ N, 127° 21′ 58″ E). One gram of soil sample was suspended in 9 mL of sterile 0.85% (wt/vol) 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 (4). 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 strain MAHUQ-72 has been deposited to the China General Microbiological Culture Collection Center with deposition number CGMCC 1.19066. The genomic DNA was extracted from the bacterial culture that was grown in R2A broth for 24 husing Solg Genomic DNA Prep kit (Solgent, Republic of Korea), according to the manufacturer's protocol. The DNA fragments were size selected using AMPure XP beads to isolate those within the desired range, typically 200–300 base pairs (bp). The library for sequencing was prepared using the Nextera XT DNA Library Prep Kit (Illumina, San Diego, USA). After 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. Reads were processed with Trimmomatic version 0.38 (5). The genome sequence was assembled by SOAPdenovo v2.04 assembler (6). The genome was annotated with the NCBI Prokaryotic Genome Annotation Pipeline version 6.3 (7–9). The quality analysis was performed by CheckM, version 1.2.4 (10). 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 bioinformatics tool (11) was used to measure *in silico* DNA-DNA hybridization (isDDH) values. To estimate the degree of pairwise relatedness

Editor Kenneth M. Stedman, Portland State University, Portland, Oregon, USA

Address correspondence to Md. Shahidul Islam, soashahed@gmail.com, or Md. Shahedur Rahman, ms.rahman@just.edu.bd.

The authors declare no conflict of interest.

Received 20 September 2025
Accepted 30 October 2025
Published 25 November 2025

TABLE 1 Genomic features of *Nocardioides* strain MAHUQ-72

Parameter	Result
Description of source	
Location	Anseong, South Korea
Time	2020
Type	Soil sample of a rice field
Sequencing summary	
Total genome length	4,732,459 bp
Coverage	83×
GC content	72.5%
Number of reads	2,982,010
Assembly report	
Contigs	23
Contig L_{50}	5
Contig N_{50}	453.6 kb
Genome length	4.7 Mb
Annotation report	
Genes (total)	4,620
CDSs (total)	4,566
CDSs (with protein)	4,555
ncRNAs	3
tRNA	48
Quality analysis	
Completeness	99.47%
Contamination	1.67%

between MAHUQ-72 and the closest type strain, BLAST-based average nucleotide identity (ANI) was calculated (12). The best matches to the strain MAHUQ-72 genome are the type strain *Nocardioides koreensis* JCM 16022 (accession GCA_039531775.1)with isDDH and ANI values of 27.6 and 84.2%, respectively.

The genome of *Nocardioides* strain MAHUQ-72 is 4,732,459 bp long with a GC content of 72.5%. The assembly contains 23 contigs with a coverage of 83×. A total of 4,620 genes were predicted, of which 4,555 are protein-coding genes. The details of genomic features are presented in Table 1.

ACKNOWLEDGMENTS

I would like to give special thanks to the World Data CentreCenter for Microorganisms (WDCM), The Institute of Microbiology of the Chinese Academy of Sciences, Beijing 100101, China for the draft genome sequencing of strain MAHUQ-72.

Md. Amdadul Huq, Conceptualization, Formal analysis, Writing – original draft, Writing – review and editing | Shahina Akter, Formal analysis, | Md. Shahidul Islam, Writing – review and editing, | Md. Shahedur Rahman, Writing – review and editing.

AUTHOR AFFILIATIONS

[1]Department of Life Sciences, College of BioNano Technology, Gachon University, Seongnam, Republic of Korea
[2]Department of Food Science and Biotechnology, Gachon University, Seongnam, Republic of Korea
[3]Department of Microbiology, Faculty of Science and Engineering, Rabindra Maitree University, Kushtia, Bangladesh
[4]Department 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. Shahidul Islam http://orcid.org/0009-0007-7572-7174
Md. Shahedur Rahman http://orcid.org/0009-0005-2028-4885

AUTHOR CONTRIBUTIONS

Md. Amdadul Huq, Conceptualization, Formal analysis, Investigation, Supervision, Writing – original draft, Writing – review and editing | Shahina Akter, Formal analysis | Md. Shahidul Islam, Supervision, Writing – review and editing | Md. Shahedur Rahman, Supervision, Writing – review and editing

DATA AVAILABILITY

The draft genome sequence of *Nocardioides* strain MAHUQ-72 was deposited in GenBank under the accession number JBQWTR000000000. The raw reads are available in SRA under the accession number SRR33604977.The version described in this paper is the first version, JBQWTR000000000.1.

REFERENCES

1. Prauser H. 1976. *Nocardioides*, a new genus of the order actinomycetales. Int J Syst Bacteriol 26:58–65. https://doi.org/10.1099/00207713-26-1-58
2. LPSN. 2025. Genus *Nocardioides*. https://lpsn.dsmz.de/genus/nocardioides.
3. Huq MA, Nam K, Rahman MS, Rahman MM, Parvez MAK, Kang KK, Akter S. 2024. *Nocardioides agri* sp. nov., isolated from garden soil. Int J Syst Evol Microbiol 74:006407. https://doi.org/10.1099/ijsem.0.006407
4. Reasoner DJ, Geldreich EE. 1985. A new medium for the enumeration and subculture of bacteria from potable water. Appl Environ Microbiol 49:1–7. https://doi.org/10.1128/aem.49.1.1-7.1985
5. 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
6. Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, et al. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. Gigascience 1:18. https://doi.org/10.1186/2047-217X-1-18
7. 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
8. 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
9. 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
10. 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
11. 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
12. Yoon SH, Ha SM, Lim JM, Kwon SJ, 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