

Draft genome sequence of *Streptomyces* sp. MNU76, isolated from Indus River sediment, reveals uncharacterized biosynthetic gene clusters

Aruna Kumari,^{1,2} Garima Suneja,^{1,2} Sonam Nain,^{1,2} Saraswati Awasthi,^{1,2} Meenakshi Sharma,¹ Rakesh Sharma^{1,2}

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT Here, we present the draft genome sequence of *Streptomyces* sp. MNU76 isolated from river sediment. The genome comprises 37 contigs, 10.5 Mb in size, and 71% GC content. The species harbors multiple predicted biosynthetic gene clusters, including several polyketide synthase pathways; the natural products they encode remain uncharacterized.

KEYWORDS genome analysis, *Streptomyces*, biosynthetic gene clusters, antifungal

Streptomyces are Gram-positive bacteria found in soil, water, and sediment, and in symbiotic relationships (1), and are known to produce numerous natural products with diverse bioactivities. Freshwater ecosystems represent a promising source of bioactive compounds with potential clinical applications (2).

We report the draft genome sequence of *Streptomyces* sp. strain MNU76 isolated from 100 g of Indus River sediment (1 g used for plating) collected at a depth of 10 cm in Leh, India (34°03'41.3" N, 77°38'21.9" E). The sample was homogenized, spread on actinomycete isolation agar, and incubated at 30°C for 6 days. The *Streptomyces* colonies were picked, grown on tryptic soy agar (pH 7.2), and incubated at 30°C for 5 days to obtain pure colonies. The genomic DNA was extracted using the DNeasy UltraClean Microbial Kit (Qiagen). The library was prepared using a Nextera XT DNA Library kit and sequenced with 2 × 300 bp reads on the MiSeq (Illumina). Paired-end raw read pairs (8,330,624) were quality-checked with FastQC v0.11.9 and trimmed using Trimmomatic v0.38 (3). The same DNA was used without size selection for library preparation with the ligation sequencing kit SQK-LSK114 (Oxford Nanopore Technologies) and sequenced on an FLO-MIN106 R9 flow cell using a MinION Mk1 (HAC base calling), yielding 213,449 Nanopore reads (mean length, 1.5 kb; N50, 2.0 kb). Trimming and adapter removal were done with Porechop v0.2.4 (4). Nanopore reads were assembled using Canu v1.8 (5) and polished with MiSeq reads via Pilon v1.8 (6). Genome completeness was assessed using checkM v1.1.3 using lineage_wf workflow (7). Phylogeny was reconstructed using TYGS v403 (8) based on the genome BLAST Distance Phylogeny method. ANI was calculated using FastANI v1.33. Genome annotation was performed using PGAP v6.8 (RefSeq) from NCBI (9). All tools were run using the default parameters. The draft genome of *Streptomyces* sp. MNU76 has a genome size of 10.5 Mb, a GC content of 71%, and a genome completeness of 99.6%. The hybrid genome assembly produced 37 contigs with an N50 value of 554 kb and genome coverage of 124x. The 16S rRNA-based phylogenetic tree indicated that *Streptomyces* sp. MNU76 is related to *Streptomyces tailanensis* TRM68348 ([GCF_008386495.1](#)), *Streptomyces akebiae* MG28 ([GCF_019599145.1](#)), and *Streptomyces gottesmaniae* DSM 3412 ([GCF_031845845.1](#)) (Fig. 1). Average nucleotide identity (ANI) comparisons with these strains yielded values of 87.6%, 92.4%, and 92%, respectively, indicating that *Streptomyces* sp. MNU76 might be

Editor Leighton Pritchard, University of Strathclyde, Glasgow, United Kingdom

Address correspondence to Rakesh Sharma, rakesh.igib@csir.res.in.

The authors declare no conflict of interest.

See the funding table on p. 3.

Received 25 November 2025

Accepted 9 March 2026

Published 15 April 2026

Copyright © 2026 Kumari et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](#).

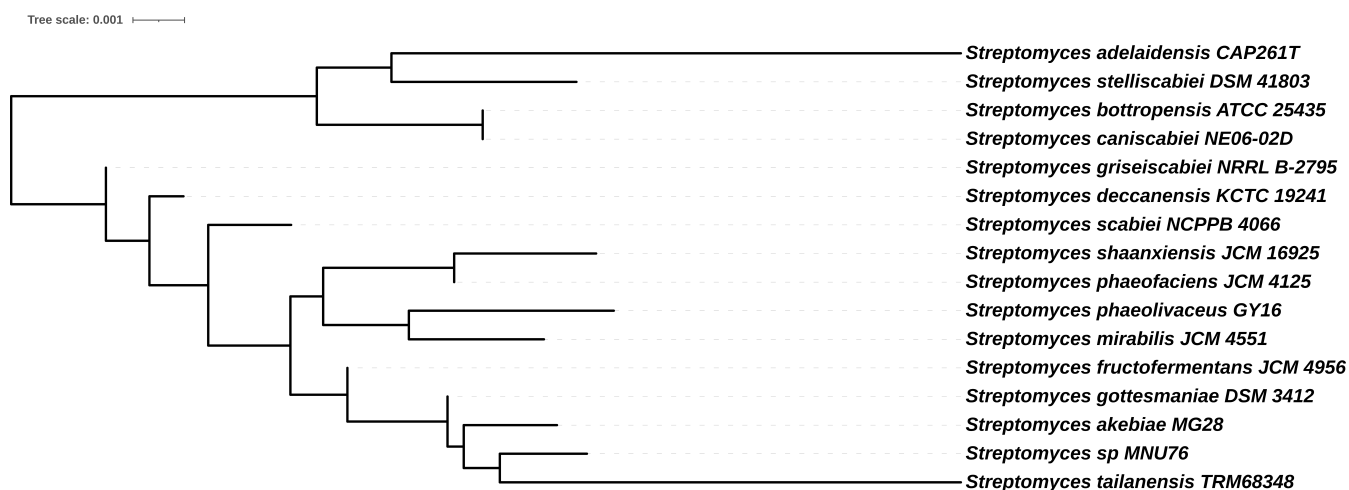


FIG 1 16S rRNA gene-based phylogenetic tree of *Streptomyces* sp. MNU76, showing the phylogenetic position of strain MNU76 with other *Streptomyces* species. The tree was generated using the TYGS server employing the genome BLAST distance phylogeny (GBDP) method on 16S rRNA sequences. The tip labels show the strain names.

a novel species. It contains 9,602 genes, comprising 8,578 coding sequences (CDSs), 72 transfer RNAs (tRNAs), 17 ribosomal RNAs (rRNAs), and three non-coding RNAs (ncRNAs) (NCBI RefSeq assembly [GCF_020982545.1](https://www.ncbi.nlm.nih.gov/assembly/GCF_020982545.1)). The antiSMASH v8.0.4 (10) analysis predicted 42 biosynthetic gene clusters (BGCs) in the genome of *Streptomyces* sp. MNU76; full, contig-resolved annotations are available via Zenodo (DOI: <https://doi.org/10.5281/zenodo.18477189>). It harbors several uncharacterized BGCs, including multiple predicted PKS-type pathways that show no similarity to known gene clusters. The diversity and richness of these clusters highlight the untapped metabolic potential of *Streptomyces* sp. MNU76. It contains several predicted PKS and NRPS-like clusters that potentially encode pharmacologically active compounds yet to be explored.

ACKNOWLEDGMENTS

We thank Dr. Souvik Maiti, Director, CSIR-IGIB, for his support and encouragement. A.K. acknowledges the Department of Biotechnology (DBT) for a Junior Research Fellowship.

This work was funded by the Council of Scientific and Industrial Research (CSIR) grant “A comprehensive approach to address antimicrobial resistance (AMR)” (MMP075202) to R.S.

A.K., Data curation, Formal Analysis, Investigation, Methodology, Validation, Software, Writing original draft | S.N., Methodology, Investigation, Software, Data curation | G.S., Methodology, Investigation, Software, Data curation | S.A., Investigation, Writing review and editing | M.S., Methodology, Investigation | R.S., Conceptualization, Data curation, Funding acquisition, Investigation, Validation, Supervision, Writing review and editing.

AUTHOR AFFILIATIONS

¹Genomics and Genome Biology, CSIR-Institute of Genomics and Integrative Biology, Council of Scientific and Industrial Research (CSIR), New Delhi, India

²Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, India

PRESENT ADDRESS

Garima Suneja, Microbiome Research Pvt Ltd, Mumbai, India

Sonam Nain, Source Bioscience, Endeavour House, Cambridge, United Kingdom

Meenakshi Sharma, Daulat Ram College, University of Delhi, Delhi, India

AUTHOR ORCID*s*

Rakesh Sharma  <http://orcid.org/0000-0003-3320-4491>

FUNDING

Funder	Grant(s)	Author(s)
Council of Scientific and Industrial Research, India	MMP075202	Rakesh Sharma

AUTHOR CONTRIBUTIONS

Aruna Kumari, Data curation, Formal analysis, Investigation, Software, Writing – original draft, Writing – review and editing | Garima Suneja, Methodology, Software, Writing – review and editing | Sonam Nain, Methodology, Software, Writing – review and editing | Saraswati Awasthi, Investigation, Writing – review and editing | Meenakshi Sharma, Formal analysis, Investigation, Methodology, Writing – review and editing | Rakesh Sharma, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Validation, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

The whole-genome shotgun project has been deposited in the National Center for Biotechnology Information (NCBI) in GenBank under accession number [GCF_020982545.1](https://doi.org/10.1093/bioinformatics/btu170). The data are available under NCBI Bioproject [PRJNA267646](https://doi.org/10.1093/bioinformatics/btu170). The raw reads have been deposited in NCBI's SRA database under accession numbers [SRR8841084](https://doi.org/10.1093/bioinformatics/btu170) (Nanopore reads) and [SRR8841085](https://doi.org/10.1093/bioinformatics/btu170) (Illumina reads). The antiSMASH annotations were uploaded to the Zenodo repository as a data set with <https://doi.org/10.5281/zenodo.18477189>.

REFERENCES

1. Barka EA, Vatsa P, Sanchez L, Gaveau-Vaillant N, Jacquard C, Meier-Kolthoff JP, Klenk H-P, Clément C, Ouhdouch Y, van Wezel GP. 2016. Taxonomy, physiology, and natural products of *Actinobacteria*. *Microbiol Mol Biol Rev* 80:1–43. <https://doi.org/10.1128/MMBR.00019-15>
2. Zothanpuia, Passari AK, Leo VV, Chandra P, Kumar B, Nayak C, Hashem A, Abd Allah EF, Alqarawi AA, Singh BP. 2018. Bioprospection of actinobacteria derived from freshwater sediments for their potential to produce antimicrobial compounds. *Microb Cell Fact* 17:68. <https://doi.org/10.1186/s12934-018-0912-0>
3. 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>
4. Wick R. 2024. Rrwick/Porechop. Available from: <https://github.com/rrwick/Porechop>
5. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. *Genome Res* 27:722–736. <https://doi.org/10.1101/gr.215087.116>
6. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. <https://doi.org/10.1371/journal.pone.0112963>
7. 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>
8. Meier-Kolthoff JP, Göker M. 2019. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat Commun* 10:2182. <https://doi.org/10.1038/s41467-019-10210-3>
9. Tatusova T, DiCuccio M, Badretin 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. Blin K, Shaw S, Vader L, Szenei J, Reitz ZL, Augustijn HE, Cediël-Becerra JDD, de Crécy-Lagard V, Koetsier RA, Williams SE, et al. 2025. antiSMASH 8.0: extended gene cluster detection capabilities and analyses of chemistry, enzymology, and regulation. *Nucleic Acids Res* 53:W32–W38. <https://doi.org/10.1093/nar/gkaf334>