

Genome sequence of *Stenotrophomonas geniculata* AS50Mn isolated from sand filters used to treat manganese-containing alkaline mine water

Hiroshi Habe,¹ Tomohiro Inaba¹

AUTHOR AFFILIATION See affiliation list on p. 2.

ABSTRACT *Stenotrophomonas geniculata* AS50Mn was isolated from sand filters used to treat manganese-containing mine water under alkaline conditions. Here, we report the complete genome sequence of this bacterium, which comprises a 4,478,846 bp circular chromosome. The genome contains 4,082 coding DNA sequences, 13 rRNA genes, and 79 tRNA genes.

KEYWORDS alkaline-tolerant bacteria, manganese-containing mine water, manganese dioxide-coated sand, rapid sand filter, *Stenotrophomonas geniculata*

Even after mine closure, continuous chemical treatment of mine water containing metals and metalloids is required (1). A pilot-scale rapid sand filter system was applied to treat manganese (Mn)-containing mine water at a mining site. When the pH of the inlet water was adjusted to 8.5–9.0, Mn removal rates reached 89.5% and 95.0%, respectively (2). During this process, *Stenotrophomonas* sp. AS50Mn was isolated from the sand filter, where it was identified as the predominant species in the microbial community (2, 3). Sand filter samples were plated on 50K medium agar (pH 8.2), containing 9.5 g/L $\text{MnSO}_4 \cdot 5\text{H}_2\text{O}$, 2.0 g/L Bacto Peptone, 0.5 g/L Bacto Yeast Extract, and 2.38 g/L (10 mM) HEPES. After several rounds of subculturing at 30°C, a single colony type was isolated, confirmed for purity, and identified as *Stenotrophomonas* (3). Strain AS50Mn demonstrated growth across Mn^{2+} concentrations ranging from 2.75 to 44 mg/L at pH 8.5 and up to 22 mg/L at pH 9.0 (3). However, the molecular basis underlying its resistance to alkaline pH and Mn^{2+} remains unclear.

A single colony of strain AS50Mn was picked from an R2A agar plate (Daigo; Shiotani M.S. Co., Ltd., Amagasaki, Japan), inoculated into 30 mL of the R2A broth liquid medium, and cultivated at 30°C for 24 h. Genomic DNA was extracted following a standard protocol (4). Genome sequencing was performed at the Taniguchi Dental Clinic Oral Microbiome Center (Kagawa, Japan) following standard library preparation workflows. For short-read sequencing, genomic libraries were prepared using the Illumina DNA Prep (M) Tagmentation Kit (Illumina, San Diego, CA, USA) according to the Kit protocol, and sequencing was performed on a MiSeq platform using the MiSeq Reagent Kit v2 (300 cycles, Illumina). For long-read sequencing, genomic libraries were prepared using the Native Barcode Sequencing Kit (SQK-NBD114-24; Oxford Nanopore Technologies, Oxford, UK) according to the Kit protocol after treatment with Short Read Eliminator XS (PacBio, Menlo Park, CA, USA) to remove <10 kb DNA, and sequenced using a PromethION 2i platform with FLO-PRO114M flow cells (Oxford Nanopore Technologies). The software used in the long-read sequence were as follows: MinKNOW v23.11.8, Bream v7.8.2, Configuration v5.8.6, Dorado v7.2.14, and MinKNOW Core v5.8.9.

A total of 500,296,458 paired-end reads (150 bp) were obtained from the Illumina platform (number of sequences: 3,285,102; N_{50} : 152), and 1,269,072,355 single-end reads

Editor Vanja Klepac-Ceraj, Wellesley College, Wellesley, Massachusetts, USA

Address correspondence to Hiroshi Habe, hiroshi.habe@aist.go.jp.

The authors declare no conflict of interest.

See the funding table on p. 2.

Received 27 October 2025

Accepted 15 February 2026

Published 27 February 2026

Copyright © 2026 Habe and Inaba. 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 Annotation statistics

	Value
Total sequence length (bp)	4,478,846
Number of sequences	1
Longest sequences (bp)	4,478,846
N_{50} (bp)	4,478,846
Gap ratio (%)	0
GC content (%)	66.5
Number of CDSs	4,082
Average protein length (aa)	327.6
Coding ratio (%)	89.6
Number of rRNAs	13
Number of tRNAs	79
Number of CRISPRs	2

were obtained from the Nanopore platform (number of sequences: 204,812; N_{50} : 6,196). Adapter and low-quality reads were trimmed using Fastp (5) v0.20.1 for Illumina data and NanoPlot (6) v1.32.1 and NanoFilt (7) v2.7.1 for Nanopore data. After quality filtering, a total of 1,221,569,407 trimmed reads were used for *de novo* assembly with Unicycler (8) v0.4.8 and Pilon (9) v1.24, applying standard evaluation strategies (10, 11). Circularity and replicon boundaries were determined using Flye (12) v2.9.1. Also, the assembly graph visualized in Bandage (13) v0.8.1 confirmed the genome to be circular. In Unicycler, the genome was rotated to start at the *dnaA* gene. Gene annotation was performed using DFAST (14) v1.6.0. Completeness was assessed using CheckM v1.1.6. Default parameters were applied for all software unless otherwise specified.

The genome consisted of 4,478,846 bp with a coverage of 111×, an average GC content of 66.5%, and 4,082 predicted coding sequences. The genome also contained 13 rRNA and 79 tRNA genes (Table 1). Based on genome BLAST distance phylogeny analysis, the genome of strain AS50Mn exhibited an average nucleotide identity of 95.9% with *S. geniculata* BR23 (accession number [CP134450](#)). According to the high similarity of whole genome sequences, strain AS50Mn was classified as *S. geniculata*. The transcriptomic analysis based on this genome may yield additional information regarding the bacterium's survival and adaptation in sand filters used to treat manganese-rich alkaline mine water.

ACKNOWLEDGMENTS

This work was supported in part by JSPS KAKENHI (Grant number 24K03111).

AUTHOR AFFILIATION

¹Environmental Management Research Institute, National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki, Japan

AUTHOR ORCID*s*

Hiroshi Habe  <http://orcid.org/0000-0001-8721-2206>

Tomohiro Inaba  <https://orcid.org/0000-0001-5330-7813>

FUNDING

Funder	Grant(s)	Author(s)
JSPS KAKENHI	24K03111	Hiroshi Habe

AUTHOR CONTRIBUTIONS

Hiroshi Habe, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Supervision, Validation, Visualization, Writing – original draft, Writing – review and editing | Tomohiro Inaba, Resources, Validation, Visualization, Writing – original draft

DATA AVAILABILITY

The genome sequence has been deposited in DDBJ/EMBL/GenBank under accession number [BAAGIK010000001](#). The raw sequence data are available under accession numbers [DRR627340](#) (Illumina) and [DRR627341](#) (Nanopore). The BioProject is registered under accession number [PRJDB19894](#).

REFERENCES

- Habe H, Sato Y, Aoyagi T, Inaba T, Hori T, Hamai T, Hayashi K, Kobayashi M, Sakata T, Sato N. 2020. Design, application, and microbiome of sulfate-reducing bioreactors for treatment of mining-influenced water. *Appl Microbiol Biotechnol* 104:6893–6903. <https://doi.org/10.1007/s00253-020-10737-2>
- Inaba T, Aizawa H, Aoyagi T, Sato Y, Hori T, Nishimura T, Habe H. 2023. Startup performance and microbial composition of a pilot-scale rapid sand filter for the treatment of manganese-containing mine water. *Chemosphere* 343:140229. <https://doi.org/10.1016/j.chemosphere.2023.140229>
- Habe H, Inaba T, Nishimura T. 2025. Characterization of dominant bacterial species on sand filters used for treating manganese-containing mine water under alkaline conditions. *Curr Microbiol* 82:584. <https://doi.org/10.1007/s00284-025-04571-w>
- Green MR, Sambrook J. 2012. *Molecular cloning: A laboratory manual*. 4th ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- 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>
- De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 34:2666–2669. <https://doi.org/10.1093/bioinformatics/bty149>
- Mann BC, Bezuidenhout JJ, Swanevelder ZH, Grobler AF. 2021. MinION 16S datasets of a commercially available microbial community enables the evaluation of DNA extractions and data analyses. *Data Brief* 36:107036. <https://doi.org/10.1016/j.dib.2021.107036>
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>
- 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>
- Laetsch DR, Blaxter ML. 2017. BlobTools: Interrogation of genome assemblies. *F1000Res* 6:1287. <https://doi.org/10.12688/f1000research.12232.1>
- Goldstein S, Beka L, Graf J, Klassen JL. 2019. Evaluation of strategies for the assembly of diverse bacterial genomes using MinION long-read sequencing. *BMC Genomics* 20:23. <https://doi.org/10.1186/s12864-018-5381-7>
- Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37:540–546. <https://doi.org/10.1038/s41587-019-0072-8>
- Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of *de novo* genome assemblies. *Bioinformatics* 31:3350–3352. <https://doi.org/10.1093/bioinformatics/btv383>
- 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>