

Draft genome sequence of *Streptococcus gallolyticus* AAM02 isolated from traditional fermented milk in Bangladesh

Md. Fakruddin,¹ Md. Abu Bakar Karim,¹ Samara Aliya Chowdhury,¹ Abeda Sultana,² Sharmin Zaman Emon,³ Md. Latiful Bari,³ Md. Arafat Al Mamun³

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT We announce the draft genome sequence of *Streptococcus gallolyticus* strain AAM02, isolated from traditional fermented milk (doi) collected in Bogura, Bangladesh (GPS: 24.85°N, 89.37°E) in 2023. The genome provides insights into its genetic adaptations to the fermented milk environment and metabolic characteristics relevant to fermentation.

KEYWORDS genomes, Bangladesh, fermented Milk

Streptococcus gallolyticus is a gram-positive, facultative anaerobic lactic acid bacterium found in traditional fermented foods. Strain AAM02 was isolated from fermented milk samples cultured on MRS agar under microaerophilic conditions at 37°C. Preliminary identification was performed by 16S rRNA gene sequencing (primers: 27F/1492R; deposited in GenBank under accession [PQ123456](#)). For whole-genome sequencing, the strain was grown on M17 agar at 37°C overnight. Genomic DNA was extracted using the Qiagen DNeasy Blood & Tissue Kit without shearing or size selection. Libraries were prepared with the Nextera DNA Flex Library Preparation Kit, and paired-end sequencing (2 × 150 bp) was performed on an Illumina NovaSeq 6000 at icddr, b Genome Center, Bangladesh. A total of 2,856,432 reads were generated, providing ~230× coverage. Raw reads were quality-checked with FastQC v0.11.9 (1) and trimmed with Trimmomatic v0.39 (2). Assembly was performed using SPAdes v4.1.0 (3), yielding 21 contigs (N50 = 263,119 bp; N90 = 59,687 bp). The genome size is 1,834,679 bp with 37.62% G+C content. Taxonomic identification was confirmed with Kraken2 v2.1.3 (4). Annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline, identifying 1,853 coding sequences (1,298 with functional assignments, 555 hypothetical proteins), 6 rRNA genes, 52 tRNA genes, and 1 tmRNA. Antibiotic resistance genes (glycopeptide) were detected using ABRicate v1.0.1 (5) with the ResFinder database, but no virulence genes were identified. Prophage screening with PHASTEST v1.0.1 (6) revealed one 33.9-kb phage-related sequence. Subsystem analysis through RAST v2.0 (7) showed 586 subsystems covering 31% of the genome.

The genome sequence of *S. gallolyticus* AAM02 provides a valuable resource for comparative genomics and understanding strain-specific adaptations in traditional Bangladeshi fermented milk production.

Editor Julie C. Dunning Hotopp, University of Maryland School of Medicine, Baltimore, Maryland, USA

Address correspondence to Md. Arafat Al Mamun, mamun1050@gmail.com.

The authors declare no conflict of interest.

Received 27 December 2025

Accepted 16 March 2026

Published 30 March 2026

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

ACKNOWLEDGMENTS

This work was supported by the Grant for Advanced Research in Education, Bangladesh Bureau of Educational Information & Statistics, Ministry of Education, Government of the People's Republic of Bangladesh (Project ID: LS20232568).

AUTHOR AFFILIATIONS

¹Department of Biochemistry and Microbiology, North South University, Dhaka, Bangladesh

²Department of Anatomy, Dhaka Medical College, Dhaka, Bangladesh

³Centre for Advanced Research in Sciences, University of Dhaka, Dhaka, Bangladesh

AUTHOR ORCIDs

Md. Fakruddin  <https://orcid.org/0000-0003-2720-9173>

Md. Arafat Al Mamun  <http://orcid.org/0000-0001-6311-8135>

AUTHOR CONTRIBUTIONS

Md. Fakruddin, Conceptualization, Formal analysis, Methodology, Visualization, Writing - original draft | Md. Abu Bakar Karim, Data curation, Formal analysis, Software, Validation, Visualization | Samara Aliya Chowdhury, Data curation, Formal analysis, Methodology, Software, Visualization | Abeda Sultana, Investigation, Methodology, Software, Validation, Visualization | Sharmin Zaman Emon, Formal analysis, Investigation, Methodology, Resources, Validation | Md. Latiful Bari, Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review and editing | Md. Arafat Al Mamun, Conceptualization, Methodology, Project administration, Supervision, Writing – review and editing

DATA AVAILABILITY

This whole-genome shotgun project has been deposited at GenBank under the accession number [JBSLYR000000000](https://www.ncbi.nlm.nih.gov/nuclseq/JBSLYR000000000). The sequenced genome is publicly available under SRA, BioProject, and BioSample numbers, which are [SRR36606348](https://www.ncbi.nlm.nih.gov/sra/SRR36606348), [PRJNA1336682](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1336682), and [SAMN52081389](https://www.ncbi.nlm.nih.gov/biosample/SAMN52081389), respectively. This genome is publicly available.

REFERENCES

1. Andrews S. 2010. Fastqc: a quality control tool for high throughput sequence data. Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom.
2. 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>
3. Pribelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. 2020. Using SPAdes *de novo* assembler. *Curr Protoc Bioinformatics* 70:e102. <https://doi.org/10.1002/cpbi.102>
4. Wood DE, Lu J, Langmead B. 2019. Improved metagenomic analysis with Kraken 2. *Genome Biol* 20:257. <https://doi.org/10.1186/s13059-019-1891-0>
5. Seemann T. 2020. Abricate: mass screening of contigs for antimicrobial resistance or virulence genes. Available from: <https://github.com/tseemann/abricate>
6. Wishart DS, Han S, Saha S, Oler E, Peters H, Grant JR, Stothard P, Gautam V. 2023. PHASTEST: faster than PHASTER, better than PHAST. *Nucleic Acids Res* 51:W443–W450. <https://doi.org/10.1093/nar/gkad382>
7. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, et al. 2008. The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. <https://doi.org/10.1186/1471-2164-9-75>