

Draft genome sequence of *Bacillus cereus* HS120, a probiotic candidate isolated from the gut of hilsa (*Tenualosa ilisha*)

Farzana Yeasmin,¹ Md. Rony Babu,¹ Promi Saha,¹ Santu Biswas,¹ Arpita Guha,¹ Tofazzal Islam,¹ Dipali Rani Gupta,¹ Md. Mahbubur Rahman¹

AUTHOR AFFILIATION See affiliation list on p. 3.

ABSTRACT *Bacillus cereus* HS120 was isolated from the stomach of a healthy hilsa (*Tenualosa ilisha*) in Nabogonga River, Bangladesh. Here, we present a 5,673,590-bp draft genome with 35.23% GC content, which encodes numerous biosynthetic gene clusters for secondary metabolites and lacks detectable virulence factors, supporting its potential as an aquaculture probiotic.

KEYWORDS *Bacillus cereus*, hilsa (*Tenualosa ilisha*), gut microbiome, probiotic, antimicrobial peptides, biosynthetic gene clusters

Hilsa (*Tenualosa ilisha*), the national fish of Bangladesh, contributes approximately 12% of the country's total fish production (1). The anadromous hilsa exhibits notable disease resilience, suggesting that its associated microbiome plays a key role in host health and pathogen defense. Recent metagenomics studies reveal the probiotic potential of the hilsa microbiome (2, 3). To explore hilsa-associated probiotic bacteria, *Bacillus cereus* HS120 was isolated from the stomach of a healthy hilsa from the Nabogonga River (23.26° N, 89.24° E), Bangladesh. Stomach tissue was aseptically processed, serially diluted, plated on Zobell agar, and incubated at 28°C for 24 h. A single colony was purified by repeated streaking. Cell-free culture filtrate of HS120 exhibited potent antagonistic activity against the fish pathogens *Aeromonas veronii*, *Streptococcus agalactiae*, and *Enterococcus faecalis* (Fig. 1) in an *in vitro* assay (4).

Genomic DNA was extracted from a 24-h culture grown in Zobell broth at 28°C with shaking at 150 rpm under aerobic conditions using the GeneJET Genomic DNA Purification Kit (Thermo Fisher Scientific, USA). DNA quality and quantity were assessed using a Qubit 4.0 fluorometer and a NanoDrop 2000 spectrophotometer, respectively. High molecular weight genomic DNA was used without intentional shearing to preserve long-read sequencing capability, and no explicit size selection step was performed. A sequencing library was prepared using the Oxford Nanopore Ligation Sequencing Kit (SQK-LSK109/114), purified with AMPure XP beads, and sequenced on a MinION Mk1B device (5). Base-called was performed using Guppy (Oxford Nanopore Technologies) with high-accuracy model.

A total of 67,954 generated raw reads were assessed using NanoPlot v1.46.1 (6), adapter trimmed with Porechop v0.2.4 (7), and filtered (Q > 10; length > 1,000 bp) with Filtrlong v0.3.1 (8). *De novo* assembly was performed with Flye v2.9.6 (9, 10) and polished with Medaka v2.1.1 (11). Assembly quality was evaluated using QUAST v5.3.0, and annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline (12). Additional genomic features were studied using RAST Server v2.0 (13), VirulenceFinder (14), and PlasmidFinder v2.0.1 (15). All analyses were performed using default parameters.

The draft genome of *B. cereus* HS120 comprises 5,673,590 bp (GC content 35.23%), assembled into six contigs (largest contig: 5,331,035 bp; smallest contig: 5,049 bp) with approximately 94 × coverage. The N50 value was 5,331,035 bp. A total of 5,843 genes

Editor Frank J. Stewart, Montana State University, Bozeman, Montana, USA

Address correspondence to Md. Mahbubur Rahman, mahbub-biotech@gau.edu.bd.

The authors declare no conflict of interest.

See the funding table on p. 3.

Received 19 February 2026

Accepted 31 March 2026

Published 30 April 2026

Copyright © 2026 Yeasmin 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/).

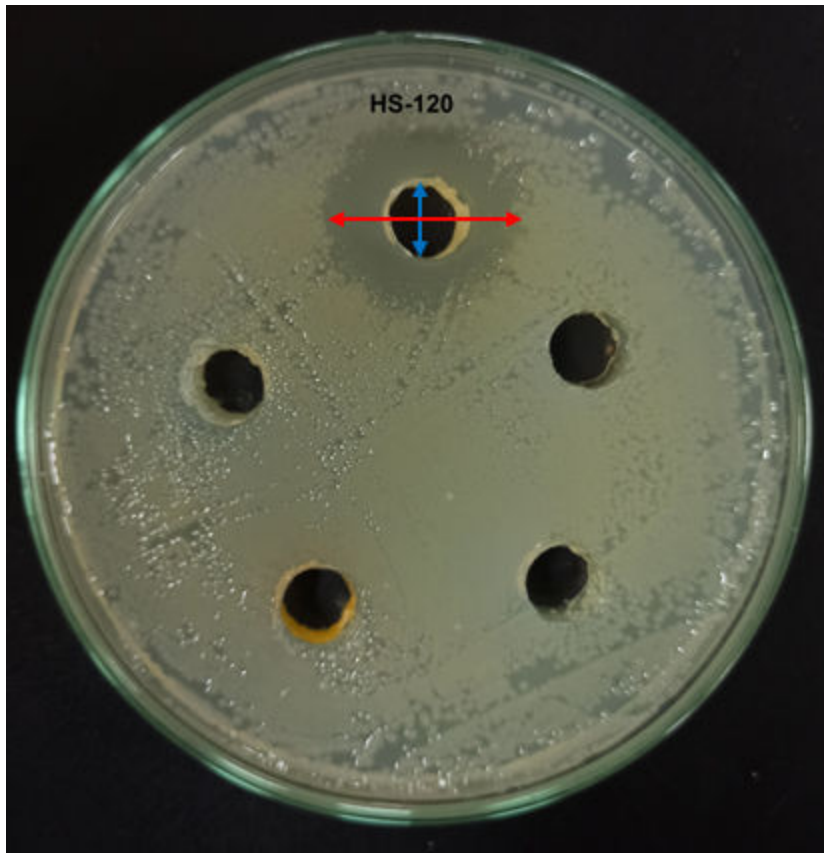


FIG 1 Inhibitory effects of cell-free culture filtrate of *B. cereus* HS120. Antimicrobial activity was evaluated using the agar well diffusion method. Fish pathogenic *Aeromonas veronii* B55 was uniformly spread onto Mueller-Hinton agar plate. Wells (blue arrow) were aseptically punched, and 50 μ L of cell-free culture filtrate of HS120 was added to one well. After 48 h of incubation, a clear zone (red arrow) surrounding the well indicated inhibition of the pathogen.

were predicted, including 5,510 protein-coding sequences and 156 RNA genes. Average Nucleotide Identity analysis (OrthoANIu) (16) revealed 98.80% similarity to *B. cereus* (BioSample ID: [SAMN44839476](#)), confirming species identity. Genomic mining identified 14 biosynthetic gene clusters via antiSMASH 8.0.4, including NRPS, RiPP, and betalactones (facilitating niche adaptation) (17), and two antimicrobial peptide loci via BAGEL4 (18). Notably, no plasmids, antibiotic resistance genes (via ResFinder 4.7.2), or virulence-associated genes (VFDB) were detected (15, 19, 20). Functional analysis highlighted genes for nutrient assimilation (metallocarboxypeptidases and aminopeptidases) (21, 22), sporulation (supporting environmental persistence) (23), TasA-mediated biofilm formation (facilitating mucosal attachment) (24), and iron-scavenging systems (bacillibactin and petrobactin) (25). This is the first genome report of a *B. cereus* strain from hilsa gut exhibiting probiotic-associated traits, supporting its safe application as a probiotic candidate for sustainable aquaculture.

ACKNOWLEDGMENTS

This work was partially supported by the National Science and Technology Fellowship provided by the Ministry of Science and Technology, Government of the People's Republic of Bangladesh.

Farzana Yeasmin, investigation, methodology, data curation, writing original draft; Md. Rony Babu, software, investigation, writing original draft; Promi Saha, methodology, writing original draft; Santu Biswas, software, investigation, writing original draft; Arpita

Guha, investigation, writing original draft; Tofazzal Islam, conceptualization, writing – review and editing; Dipali Rani Gupta, supervision, writing – review and editing; Mahbubur Rahman, conceptualization, project administration, resources, supervision, validation, writing – review and editing.

AUTHOR AFFILIATION

¹Institute of Biotechnology and Genetic Engineering, Gazipur Agricultural University, Gazipur, Bangladesh

AUTHOR ORCID*s*

Farzana Yeasmin  <http://orcid.org/0009-0009-4348-0538>

Md. Mahbubur Rahman  <http://orcid.org/0000-0003-1181-6817>

FUNDING

Funder	Grant(s)	Author(s)
Ministry of Science and Technology, Government of the People's Republic of Bangladesh	673	Farzana Yeasmin

AUTHOR CONTRIBUTIONS

Farzana Yeasmin, Data curation, Formal analysis, Investigation, Methodology, Software, Writing – original draft | Md. Rony Babu, Formal analysis, Methodology, Software, Validation, Writing – original draft | Promi Saha, Data curation, Methodology, Writing – original draft | Santu Biswas, Investigation, Methodology, Software, Writing – original draft | Arpita Guha, Methodology, Writing – original draft | Tofazzal Islam, Conceptualization, Writing – review and editing | Dipali Rani Gupta, Writing – review and editing | Md. Mahbubur Rahman, Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – review and editing

DATA AVAILABILITY

The complete genome sequence of *Bacillus cereus* HS120 has been deposited in GenBank under the accession number [JBTXPM000000000](https://www.ncbi.nlm.nih.gov/nuclseq/JBTXPM000000000). The raw sequencing reads have been deposited in the NCBI Sequence Read Archive (SRA) under the accession number [SRR37258482](https://www.ncbi.nlm.nih.gov/sra/SRR37258482). The data can be found under BioSample accession number [SAMN53301936](https://www.ncbi.nlm.nih.gov/biosample/SAMN53301936) and BioProject accession number [PRJNA1366063](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1366063).

REFERENCES

1. Foyals MJ, Momtaz F, Robiul Kawser AQM, Chaklader MR, Siddik MAB, Lamichhane B, Tay ACY, Rahman MM, Fotedar R. 2019. Microbiome patterns reveal the transmission of pathogenic bacteria in hilsa fish (*Tenualosa ilisha*) marketed for human consumption in Bangladesh. *J Appl Microbiol* 126:1879–1890. <https://doi.org/10.1111/jam.14257>
2. Islam T, Hoque MN. 2023. Gut and flesh microbiome sequencing of the Bangladesh national fish hilsa (*Tenualosa ilisha*). *Microbiol Resour Announc* 12:e0044823. <https://doi.org/10.1128/MRA.00448-23>
3. Kawser A, Hoque MN, Rahman MS, Sakif TI, Coffey TJ, Islam T. 2024. Unveiling the gut bacteriome diversity and distribution in the national fish hilsa (*Tenualosa ilisha*) of Bangladesh. *PLoS One* 19:e0303047. <https://doi.org/10.1371/journal.pone.0303047>
4. Rahman MM, Paul SI, Rahman A, Haque MS, Ador MAA, Foyals MJ, Islam MT, Rahman MM. 2022. Suppression of streptococcosis and modulation of the gut bacteriome in Nile Tilapia (*Oreochromis niloticus*) by the marine sediment bacteria *Bacillus haynesii* and *Advenella mimigardefordensis*. *Microbiol Spectr* 10:e0254222. <https://doi.org/10.1128/spectrum.02542-22>
5. Lu H, Giordano F, Ning Z. 2016. Oxford Nanopore MinION sequencing and genome assembly. *Genomics Proteomics Bioinformatics* 14:265–279. <https://doi.org/10.1016/j.gpb.2016.05.004>
6. 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>
7. Rrwick. 2017. GitHub - rrrwick/Porechop: adapter trimmer for Oxford Nanopore reads. <https://github.com/rrrwick/Porechop>.
8. Rrwick. 2017. GitHub - rrrwick/Filtlong: quality filtering tool for long reads. GitHub. <https://github.com/rrrwick/Filtlong>.
9. Lin Y, Yuan J, Kolmogorov M, Shen MW, Chaisson M, Pevzner PA. 2016. Assembly of long error-prone reads using de Bruijn graphs. *Proc Natl Acad Sci USA* 113:E8396–E8405. <https://doi.org/10.1073/pnas.1604560113>
10. Mikolmogorov. 2021. GitHub - mikolmogorov/Flye: de novo assembler for single molecule sequencing reads using repeat graphs. GitHub. <https://github.com/mikolmogorov/Flye>.

11. Nanoporetech. n.d. 2020. GitHub - nanoporetech/medaka: sequence correction provided by ONT Research. GitHub. <https://github.com/nanoporetech/medaka>.
12. Mikheenko A, Prjibelski A, Saveliev V, Antipov D, Gurevich A. 2018. Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics* 34:i142–i150. <https://doi.org/10.1093/bioinformatics/bty266>
13. 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>
14. Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS, Nielsen EM, Aarestrup FM. 2014. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. *J Clin Microbiol* 52:1501–1510. <https://doi.org/10.1128/JCM.03617-13>
15. Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, Villa L, Møller Aarestrup F, Hasman H. 2014. *In silico* detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother* 58:3895–3903. <https://doi.org/10.1128/AAC.02412-14>
16. Lee I, Ouk Kim Y, Park S-C, Chun J. 2016. OrthoANI: an improved algorithm and software for calculating average nucleotide identity. *Int J Syst Evol Microbiol* 66:1100–1103. <https://doi.org/10.1099/ijsem.0.000760>
17. Blin K, Medema MH, Kottmann R, Lee SY, Weber T. 2017. The antiSMASH database, a comprehensive database of microbial secondary metabolite biosynthetic gene clusters. *Nucleic Acids Res* 45:D555–D559. <https://doi.org/10.1093/nar/gkw960>
18. van Heel AJ, de Jong A, Song C, Viel JH, Kok J, Kuipers OP. 2018. BAGEL4: a user-friendly web server to thoroughly mine RiPPs and bacteriocins. *Nucleic Acids Res* 46:W278–W281. <https://doi.org/10.1093/nar/gky383>
19. Florensa AF, Kaas RS, Clausen P, Aytan-Aktug D, Aarestrup FM. 2022. ResFinder - an open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes. *Microb Genom* 8:000748. <https://doi.org/10.1099/mgen.0.000748>
20. Chen L. 2004. VFDB: a reference database for bacterial virulence factors. *Nucleic Acids Res* 33:D325–D328. <https://doi.org/10.1093/nar/gki008>
21. Arolas JL, Vendrell J, Aviles FX, Fricker LD. 2007. Metalloprotease: emerging drug targets in biomedicine. *Curr Pharm Des* 13:349–366. <https://doi.org/10.2174/138161207780162980>
22. Jankiewicz U, Bielawski W. 2003. The properties and functions of bacterial aminopeptidases. *Acta microbio Pol* 52:217–231.
23. de Vries YP, Hornstra LM, de Vos WM, Abee T. 2004. Growth and sporulation of *Bacillus cereus* ATCC 14579 under defined conditions: temporal expression of genes for key sigma factors. *Appl Environ Microbiol* 70:2514–2519. <https://doi.org/10.1128/AEM.70.4.2514-2519.2004>
24. Caro-Astorga J, Pérez-García A, de Vicente A, Romero D. 2014. A genomic region involved in the formation of adhesin fibers in *Bacillus cereus* biofilms. *Front Microbiol* 5:745. <https://doi.org/10.3389/fmicb.2014.00745>
25. Lee JY, Passalacqua KD, Hanna PC, Sherman DH. 2011. Regulation of petrobactin and bacillibactin biosynthesis in *Bacillus anthracis* under iron and oxygen variation. *PLoS One* 6:e20777. <https://doi.org/10.1371/journal.pone.0020777>