

Whole-genome sequence of *Bacillus subtilis* TP111, a potential fish probiotic that prevents motile *Aeromonas* septicemia in Nile tilapia (*Oreochromis niloticus*)

Sulav Indra Paul,¹ Ashikur Rahman,¹ Md. Javed Foysal,² Md. Mahbubur Rahman¹

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT We report the genome of fish probiotic *Bacillus subtilis* TP111 strain isolated from the gut of a healthy Nile tilapia (*Oreochromis niloticus*) in Bangladesh. TP111 has a genome size of 4,174,638 bp, 43.48% guanine-cytosine, 243.0× genome coverage with 4,224 potential coding sequences, and 10 predicted secondary metabolite biosynthetic gene clusters.

KEYWORDS *Bacillus subtilis*, antimicrobial peptides, fish probiotic, motile *Aeromonas* septicemia

Bacillus subtilis is recognized as a probiotic bacterium frequently studied for its beneficial properties and safety profile (1, 2). Here, we report the whole-genome sequence of a promising fish probiotic strain, TP111, which exhibits *in vitro* antimicrobial activity against fish pathogenic *Aeromonas veronii* and suppresses motile *Aeromonas* septicemia in Nile tilapia (*Oreochromis niloticus*) (3).

To isolate *Bacillus subtilis* TP111, the abdomen of a healthy Nile tilapia was cut aseptically, and the gut was taken out. One gram of homogenates of the intestinal segments was serially diluted and spread onto de Man, Rogosa, and Sharpe (MRS) agar plates and incubated at 28°C for 48 h. TP111 was picked from the growing colonies on the MRS plate (3). To isolate high-quality genomic DNA, a single colony of TP111 was inoculated in MRS broth and incubated at 28°C for 48 h. Then, the DNA was extracted using a GeneJET genomic DNA purification kit (Thermo Fisher Scientific, USA) according to the manufacturer's instructions. Extracted DNA was quantified using a NanoDrop spectrophotometer (Thermo Fisher Scientific). A paired-end DNA library was prepared using a Nextera XT library prep kit (Illumina, San Diego, CA, USA) according to the manufacturer's instructions (4). Genome sequencing (600 cycles) was carried out using the Illumina MiSeq benchtop sequencer (Illumina), yielding a total of 4,339,952 paired-end reads with 1,013,366,396 bases. Quality filtering was done using PRINSEQ v.0.20.3 (5), and Trimmomatic v.0.38 (6) was used for trimming low-quality sequences. The *de novo* assembly was conducted using SPAdes v.3.9.0 (7) followed by gene prediction and annotation using the National Center for Biotechnology Information Prokaryotic Genome Annotation Pipeline (PGAP) (<https://www.ncbi.nlm.nih.gov/refseq/annotation/prok/>) (8). Secondary metabolite biosynthetic gene clusters were identified using antiSMASH v.6.0 (9). Probiotic safety-associated genes were checked using ResFinder v.4.1 (10) and PathogenFinder v.1.1 (11). Default parameters were used for all software unless otherwise noted.

The *de novo* assembly resulted in an estimated chromosome size of 4,174,638 bp (26 contigs), with 43.48% guanine-cytosine content from 4,339,952 paired-end reads and a total of 1,013,366,396 bases sequenced, providing 243× genome coverage. The genome contains 4,224 coding sequences and 106 RNA genes as predicted by PGAP (81 tRNA,

Editor Irene L. G. Newton, Indiana University, Bloomington, Indiana, USA

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

The authors declare no conflict of interest.

See the funding table on p. 3.

Received 20 December 2024

Accepted 22 March 2025

Published 21 April 2025

Copyright © 2025 Paul 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/).

TABLE 1 Presence of secondary metabolite biosynthetic gene clusters in the genome sequence of *Bacillus subtilis* TP111, as predicted by antiSMASH

Region	Type	From	To	Most similar known cluster		Similarity
Scaffold 2						
Region 2.1	T3PKS	94,603	135,700	1-Carbapen-2-em-3-carboxylic acid	Other	16
Scaffold 4						
Region 4.1	NRPS	191,707	257,095	Surfactin	NRP: lipopeptide	82
Region 4.2	Sactipeptide and ranthipeptide	386,206	409,159	Sporulation killing factor	RiPP: head-to-tailcyclized peptide	100
Scaffold 5						
Region 5.1	Sactipeptide	8,743	30,354	Subtilisin A Pulcherriminic acid	RiPP:thiopeptide	100
Region 5.2	CDPS	241,961	262,707		Other	100
Scaffold 6						
Region 6.1	TransAT-PKS, PKS-like, T3PKS, and NRPS	270,063	384,234	Bacillaene	Polyketide + NRP	100
Scaffold 7						
Region 7.1	Eipeptide	85,458	107,156	Thailanstatin ABacilysin	NRP + polyketide	10
Region 7.2	Other	325,496	361,282		Other	100
Scaffold 8						
Region 8.1	NRP-metallophore, NRPs	81,682	133,459	Bacillibactin	NRP	100
Scaffold 10						
Region 10.1	NRPs, betalactone	47,522	125,281	Fengycin	NRP	100

20 rRNA, and five non-coding RNA genes). The N_{50} and L_{50} values of the assembly were 389,041 and 5, respectively. The largest and smallest contigs were 502,094 and 669 bp, respectively. No remarkable antibiotic-resistant genes except *aadk*, *mph(K)*, and *tet* were identified in the genome using ResFinder v.4.1 (10). PathogenFinder v.1.1 (11) predicted TP111 as a non-human pathogen (matched pathogenic families: 0, matched non-pathogenic families: 275). RAST v.2.0 (12) predicted 337 subsystems and 1,705 protein-coding genes involved in the putative functional categories of a potential probiotic bacterium. antiSMASH v.6.0 (9) predicted 10 secondary metabolite biosynthetic gene clusters (Table 1). The presented genome information will assist further specific studies of this strain to exploit its probiotic potential.

ACKNOWLEDGMENTS

We acknowledge the Research Management Wing of Bangabandhu Sheikh Mujibur Rahman Agricultural University (Gazipur, Bangladesh) for providing research grants for the research project Development of Native Probiotics, Herbal Extracts, and Antimicrobial Agents for Sustainable Management of Major Fish Diseases of Bangladesh.

AUTHOR AFFILIATIONS

¹Institute of Biotechnology and Genetic Engineering, Bangabandhu Sheikh Mujibur Rahman Agricultural University, Gazipur, Bangladesh

²Department of Genetic Engineering and Biotechnology, Shahjalal University of Science and Technology, Sylhet, Bangladesh

AUTHOR ORCIDs

Sulav Indra Paul  <http://orcid.org/0000-0001-8331-8011>

Ashikur Rahman  <http://orcid.org/0000-0003-4787-4920>

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

FUNDING

Funder	Grant(s)	Author(s)
Bangabandhu Sheikh Mujibur Rahman Agricultural University		Md. Mahbubur Rahman

AUTHOR CONTRIBUTIONS

Sulav Indra Paul, Data curation, Formal analysis, Methodology, Writing – original draft | Ashikur Rahman, Investigation, Methodology | Md. Javed Foysal, Data curation, Formal analysis, Investigation, Validation, Writing – review and editing | Md. Mahbubur Rahman, Conceptualization, Data curation, Funding acquisition, Investigation, Project administration, Resources, Supervision, Validation, Writing – review and editing

DATA AVAILABILITY

The whole-genome shotgun project of *B. subtilis* strain TP111 has been deposited at GenBank under assembly accession number [JAXIVG000000000](https://www.ncbi.nlm.nih.gov/assembly/JAXIVG000000000/). Raw sequence reads are available under SRA accession number [SRX22722889](https://www.ncbi.nlm.nih.gov/sra/SRX22722889), BioProject accession number [PRJNA1048088](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1048088), and BioSample accession number [SAMN38606862](https://www.ncbi.nlm.nih.gov/biosample/SAMN38606862).

ETHICS APPROVAL

Prior permission was taken from the Institute of Biotechnology and Genetic Engineering (IBGE) ethical review committee for the animal experiments (approval no. IBGE-ERC-008).

REFERENCES

1. Elshaghabe FMF, Rokana N, Gulhane RD, Sharma C, Panwar H. 2017. *Bacillus* as potential probiotics: status, concerns, and future perspectives. *Front Microbiol* 8:1490. <https://doi.org/10.3389/fmicb.2017.01490>
2. Paul SI, Rahman MdM, Salam MA, Khan MdAR, Islam MdT. 2021. Identification of marine sponge-associated bacteria of the Saint Martin's island of the Bay of Bengal emphasizing on the prevention of motile *Aeromonas septicemia* in *Labeo rohita*. *Aquaculture* 545:737156. <https://doi.org/10.1016/j.aquaculture.2021.737156>
3. Rahman A. 2019. Isolation and identification of indigenous fish gut bacteria and their disease suppression effects in tilapia, Bangabandhu Sheikh Mujibur Rahman Agricultural University, Gazipur, Bangladesh
4. Illumina. 2019. Nextera XT DNA library prep reference guide. Illumina, San Diego, CA.
5. Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27:863–864. <https://doi.org/10.1093/bioinformatics/btr026>
6. 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>
7. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>
8. Tatusova T, DiCuccio M, Badretdin A, Chetvermin 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>
9. Blin K, Wolf T, Chevrette MG, Lu X, Schwalen CJ, Kautsar SA, Suarez Duran HG, Santos ELC, Kim HU, Nave M, Dickschat JS, Mitchell DA, Shelest E, Breitling R, Takano E, Lee SY, Weber T, Medema MH. 2017. antiSMASH 4.0—improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res* 45:W36–W41. <https://doi.org/10.1093/nar/gkx319>
10. Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Cattoir V, Philippon A, Allesoe RL, Rebelo AR, Florensa AF, et al. 2020. ResFinder 4.0 for predictions of phenotypes from genotypes. *J Antimicrob Chemother* 75:3491–3500. <https://doi.org/10.1093/jac/dkaa345>
11. Cosentino S, Voldby Larsen M, Møller Aarestrup F, Lund O. 2013. PathogenFinder—distinguishing friend from foe using bacterial whole genome sequence data. *PLoS ONE* 8:e77302. <https://doi.org/10.1371/journal.pone.0077302>
12. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). *Nucleic Acids Res* 42:D206–14. <https://doi.org/10.1093/nar/gkt1226>