

# UC Davis

## UC Davis Previously Published Works

### Title

Draft genome sequence of antibiotic-resistant *Shigella flexneri* MTR\_GR\_V146 strain isolated from a tomato (*Solanum lycopersicum*) sample collected from a peri-urban area of Bangladesh.

### Permalink

<https://escholarship.org/uc/item/9xr4p5nq>

### Journal

Microbiology Resource Announcements, 13(4)

### Authors

Islam, Md Saiful

Pramanik, Pritom

Rana, Md

et al.

### Publication Date

2024-04-11

### DOI

10.1128/mra.00099-24

Peer reviewed

# Draft genome sequence of antibiotic-resistant *Shigella flexneri* MTR\_GR\_V146 strain isolated from a tomato (*Solanum lycopersicum*) sample collected from a peri-urban area of Bangladesh

Md. Saiful Islam,<sup>1,2</sup> Pritom Kumar Pramanik,<sup>1</sup> Md. Liton Rana,<sup>1</sup> Md. Ashek Ullah,<sup>1</sup> Fahim Haque Nelay,<sup>1</sup> Srinivasan Ramasamy,<sup>3</sup> Pepijn Schreinemachers,<sup>4</sup> Ricardo Oliva,<sup>3</sup> Md. Tanvir Rahman<sup>1</sup>

**AUTHOR AFFILIATIONS** See affiliation list on p. 2.

**ABSTRACT** This study announces the genome sequence of the *Shigella flexneri* MTR\_GR\_V146 strain isolated from a tomato (*Solanum lycopersicum*) sample in Bangladesh. This strain has a 4,624,521 bp genome length (coverage: 73.07×), 2 CRISPR arrays, 1 plasmid, 52 predicted antibiotic resistance genes, and 53 virulence factor genes.

**KEYWORDS** vegetables, gardening systems, *Shigella flexneri*, draft genome, ARGs, VFGs, public health, Bangladesh

The global emergence of foodborne pathogens has become a major concern for public health (1). The presence of antibiotic-resistant *Shigella* spp. in vegetables presents a public health risk as it can transfer to the human population through the food supply chain.

From September 2022 to March 2023, samples of fresh tomatoes (*Solanum lycopersicum*) were collected from various gardening systems in the Gazipur district of Bangladesh (24.0958°N, 90.4125°E) and subsequently transported to the laboratory at Bangladesh Agricultural University (24.7245°N, 90.4372°E). As previously described (2), the samples were aseptically cut, measured (50 g), and introduced into a sterile polyethylene stomacher bag with 200 mL of buffered peptone water. Samples were macerated for 5 minutes at 230 rpm using a Stomacher 400 circulator (Seward Ltd., London, UK). The processed samples were then incubated at 37°C for 24 h, spread on Salmonella-Shigella agar (HiMedia, Mumbai, Maharashtra, India) plates, and incubated at 37°C overnight. The resulting large, circular, convex, and transparent colonies were subjected to Gram staining and biochemical tests (oxidase, urease, carbohydrate fermentation test or mannitol, and H<sub>2</sub>S tests) to isolate *Shigella* spp. (3). In this study, the MTR\_GR\_V146 strain was chosen and incubated in nutrient broth (HiMedia, Mumbai, Maharashtra, India) at 37°C overnight. DNA was subsequently extracted from the cultured broth using a DNeasy Blood and Tissue Kit (QIAGEN, Hilden, Germany). The DNA library was prepared using the Nextera DNA Flex Library Prep Kit (Illumina, San Diego, CA, USA). Genome sequencing was carried out on the Illumina NextSeq2000 platform, which generated paired-end reads with a length of 2 × 150 bp. The genome assembly was conducted using Unicycler.v0.4.9 (4), following a preliminary step of trimming the raw paired-end reads (pre-trimming: 2,806,842; post-trimming: 2,753,822) with Trimmomatic.v0.39 (5) (leading: 20, sliding window: 4:20:20, trailing: 20, and minlen = 36) to remove Illumina adapters and phiX reads from the data set. Quality assessment was performed using FastQC.v0.11.7 (6). The SpeciesFinder.v2.0 (7) was used to identify *Shigella flexneri*. The annotation of the genome was done using PGAP.v6.6 (8). In

**Editor** David Rasko, University of Maryland School of Medicine, Baltimore, Maryland, United States

Address correspondence to Md. Tanvir Rahman, tanvirahman@bau.edu.bd.

The authors declare no conflict of interest.

See the funding table on p. 2.

**Received** 1 February 2024

**Accepted** 11 February 2024

**Published** 27 February 2024

Copyright © 2024 Islam 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/).

the assembled genome, pathogenicity index by PathogenFinder.v1.1 (9); CRISPR arrays and prophages by CRISPRimmunity (10); plasmids by PlasmidFinder.v2.1 (11); antibiotic resistance genes (ARGs) by CARD.v3.2.4 with the Resistance Gene Identifier(RGI) main (12); and virulence factor genes (VFGs) by VFDB with VFAnalyzer (13). Default parameters were used for all software unless otherwise specified.

In SpeciesFinder, the isolate was identified as *S. flexneri*. The assembled *S. flexneri* MTR\_GR\_V146 strain had 89 contigs, a GC content of 50.8%, and eight L50 contigs with a value of 182,395 bp for N50. The overall genome size was 4,624,521 bp with a coverage of 73.07×. The genome featured 4,514 genes, 4,423 CDS, 91 RNA genes (tRNAs – 76, rRNAs – 5, and non-coding RNAs – 10), 110 pseudogenes, 2 CRISPR arrays with 10 genes (*csa3*, *cas2*, *cas1*, *cas6e*, *cas5*, *cas7*, *cse2gr11*, *cas8e*, *cas3*, and *WYL*), 10 prophages, and 1 plasmid [IncFIB(AP001918)]. The genome exhibited a pathogenicity index of 0.943. Additionally, 52 predicted ARGs and 53 predicted VFGs were identified in the genome.

### ACKNOWLEDGMENTS

This work was conducted as part of the CGIAR Research Initiative on Resilient Cities Through Sustainable Urban and Peri-urban Agri-food Systems and is supported by contributors to the CGIAR Trust Fund (<https://www.cgiar.org/funders>).

### AUTHOR AFFILIATIONS

<sup>1</sup>Department of Microbiology and Hygiene, Faculty of Veterinary Science, Bangladesh Agricultural University, Mymensingh, Bangladesh

<sup>2</sup>Department of Animal Sciences, University of California—Davis, Davis, California, USA

<sup>3</sup>World Vegetable Center, Shanhua, Tainan, Taiwan

<sup>4</sup>World Vegetable Center – East and Southeast Asia, Bangkok, Thailand

### AUTHOR ORCIDs

Md. Saiful Islam  <https://orcid.org/0000-0002-6870-4595>

Md. Liton Rana  <https://orcid.org/0000-0003-1400-1250>

Md. Tanvir Rahman  <http://orcid.org/0000-0001-5432-480X>

### FUNDING

Funder	Grant(s)	Author(s)
Consortium of International Agricultural Research Centers (CGIAR)		Md. Tanvir Rahman

### AUTHOR CONTRIBUTIONS

Md. Saiful Islam, data curation, formal analysis, investigation, methodology, software, writing - original draft, Writing – review and editing | Pritom Kumar Pramanik, data curation, investigation, methodology | Md. Liton Rana, investigation | Md. Ashek Ullah, investigation | Fahim Haque Neloy, investigation | Srinivasan Ramasamy, Conceptualization, Validation, Writing – review and editing | Pepijn Schreinemachers, Conceptualization, Validation, Writing – review and editing | Ricardo Oliva, Conceptualization, Project administration, Supervision, Validation, Writing – review and editing | Md. Tanvir Rahman, Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Validation, Writing – review and editing

### DATA AVAILABILITY

The WGS shotgun analysis of *S. flexneri* MTR\_GR\_V146 was deposited to GenBank under the accession number [JAVTVS000000000](https://www.ncbi.nlm.nih.gov/nuclseq/JAVTVS000000000). The relevant data, including the raw reads, were submitted with BioProject accession number [PRJNA1019910](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1019910), BioSample accession

number [SAMN37503358](#), and SRA accession number [SRR26156559](#). In this version, the specific version being referred to is identified as [JAVTVS000000000.1](#).

## REFERENCES

1. Pakbin B, Amani Z, Allahyari S, Mousavi S, Mahmoudi R, Brück WM, Peymani A. 2021. Genetic diversity and antibiotic resistance of *Shigella* spp. isolates from food products. *Food Sci Nutr* 9:6362–6371. <https://doi.org/10.1002/fsn3.2603>
2. Islam MS, Pramanik PK, Rana ML, Ramasamy S, Schreinemachers P, Oliva R, Rahman MT. 2024. Draft genome sequences of five multidrug-resistant *Escherichia coli* strains isolated from vegetable samples in Bangladesh. *Microbiol Resour Announc* 13:e00982-23. <https://doi.org/10.1128/mra.00982-23>
3. Karmoker J, Islam S, Liton Rana M, Ullah A, Neloy FH, Oishy NM, Pramanik PK, Siddique MP, Saha S, Rahman T. 2023. Molecular detection and multidrug resistance of *Shigella* spp. Isolated from wild waterfowl and migratory birds in Bangladesh. *Veterinary Medicine International*. <https://doi.org/10.1155/2023/5374216>
4. 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>
5. 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>
6. Andrews S. 2010. Online. FastQC: a quality control tool for high throughput sequence data. Available from: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
7. Clausen P, Aarestrup FM, Lund O. 2018. Rapid and precise alignment of raw reads against redundant databases with KMA. *BMC Bioinformatics* 19:307. <https://doi.org/10.1186/s12859-018-2336-6>
8. Li W, O'Neill KR, Haft DH, DiCuccio M, Chetvernin V, Badretdin A, Coulouris G, Chitsaz F, Derbyshire MK, Durkin AS, Gonzales NR, Gwadz M, Lanczycki CJ, Song JS, Thanki N, Wang J, Yamashita RA, Yang M, Zheng C, Marchler-Bauer A, Thibaud-Nissen F. 2021. RefSeq: expanding the prokaryotic genome annotation pipeline reach with protein family model curation. *Nucleic Acids Res* 49:D1020–D1028. <https://doi.org/10.1093/nar/gkaa1105>
9. 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>
10. Zhou F, Yu X, Gan R, Ren K, Chen C, Ren C, Cui M, Liu Y, Gao Y, Wang S, Yin M, Huang T, Huang Z, Zhang F. 2023. CRISPRimmunity: an interactive web server for CRISPR-associated important molecular events and modulators used in geNome editing Tool identifying. *Nucleic Acids Res* 51:W93–W107. <https://doi.org/10.1093/nar/gkad425>
11. 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>
12. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, Huynh W, Nguyen A-LV, Cheng AA, Liu S, et al. 2020. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res* 48:D517–D525. <https://doi.org/10.1093/nar/gkz935>
13. Liu B, Zheng D, Zhou S, Chen L, Yang J. 2022. VFDB 2022: a general classification scheme for bacterial virulence factors. *Nucleic Acids Res* 50:D912–D917. <https://doi.org/10.1093/nar/gkab1107>