

Complete genome sequence of *Xanthomonas campestris* pv. *campestris* strain NSTU_AG-1 causing black rot disease isolated from infected cabbage leaves in Bangladesh

Arnab Goswami,¹ Maksudur R. Nayem,² Pijush K. Jhan,¹ Amena Khatun,¹ A.B.Z. N. Rahman,³ Mahmudul Hasan,³ Swagato Dutta,³ Md T. Islam,⁴ Jong-In Park,⁵ Mehede H. Rubel¹

AUTHOR AFFILIATIONS See affiliation list on p. 3.

ABSTRACT *Xanthomonas campestris* pv. *campestris* (NSTU_AG-1), the pathogen of black rot disease, was isolated from cabbage in Bangladesh. The genome size was 5,120,760 bp with guanine + cytosine (G + C) content of 64.98% and had no plasmid. The complete genome sequence data of this Xcc will help to understand the molecular mechanisms of the pathogen.

KEYWORDS genome sequence, black rot, *Xanthomonas campestris* pv. *campestris*

Xanthomonas campestris pv. *campestris* (Xcc) is a highly destructive bacterial pathogen that causes black rot disease in crucifer crops, leading to significant yield losses (>50%) worldwide, particularly in cabbage (1, 2). Here, we present the complete genome sequence of Xcc strain NSTU AG-1 isolated from infected leaves of a cabbage plant (Green Star) showing typical V-shaped chlorotic lesions (3). The leaves were collected during field sampling in Churamonkathi Union (23.22135°N, 89.16882°E), Jessore, Bangladesh. The infected leaf samples were surface sterilized with 3% sodium hypochlorite (NaOCl) for 1 min, followed by several washes with sterile distilled water. Then, only the infected portions of leaf samples containing lesions were cut into small pieces (around 2–3 mm), placed on a King's medium B base agar plate, and incubated at 30°C for 24–48 h, followed by isolation and purification through subcultures maintaining the same media and conditions (4, 5). The light-yellow colony was picked and inoculated in 10 mL of nutrient yeast broth (NYB) ([2.5% nutrient broth + 0.5% yeast extract] wt./vol.) at 160 rpm and 30°C for DNA extraction (6). Subsequently, the liquid culture broth was sent to DNA Solution, Ltd., Dhaka, Bangladesh, maintaining 4°C, where the whole-genome sequencing was performed. DNA was extracted from bacteria using a QIAamp DNA Mini Kit (QIAGEN, Germany) and quantified with a Qubit 4 fluorometer (Thermo Fisher Scientific, USA). Using the native barcoding kit (SQK-NBD-114.24), the library was prepared from extracted bacterial DNA without shearing or size selection. The DNA was ligated with ONT barcode and adapter and sequenced in the Oxford Nanopore MinION Mk1C Platform utilizing the FLO-MIN114 (R10.4.1) flow cell (Oxford Nanopore Technologies, UK). Data were collected using MINKNOW.1.11.5 (Oxford Nanopore Technologies, UK). Using Guppy ver. 6.3.2 (Oxford Nanopore Technologies) in high accuracy mode, MinION Mk1C sequencing data (FAST 5 format) were basecalled to generate pass reads (FASTQ format) with a mean quality score > 9. The number of raw reads was 17,260, and the read length of N_{50} was 14,077. The adaptor and barcode sequences were trimmed using Porechop v0.2.4 (<https://github.com/rrwick/Porechop>), while Nanofilt v2.8.0 was used to trim low-quality bases ($q < 10$) (7). The quality-trimmed reads were assembled into one contig using Flye Version 2.9.5 through *de novo* assembly method (8). Flye provided a circular assembled genome where the genome length was 5,120,760 bp

Editor Leighton Pritchard, University of Strathclyde, Glasgow, United Kingdom

Address correspondence to Mehede H. Rubel, mehede@nstu.edu.bd.

The authors declare no conflict of interest.

See the funding table on p. 3.

Received 20 December 2025

Accepted 20 April 2026

Published 18 May 2026

Copyright © 2026 Goswami 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 Genomic features of *Xanthomonas campestris* pv. *campestris* strain NSTU_AG-1

Features	Value
Genome length	5,120,760 bp
Contigs	1
G + C content	64.98%
Plasmids	0
<i>N</i> ₅₀	5,120,760
Genome completeness	98.25%
Contamination	1.11%
tRNA	54
rRNA	6
Total coding sequences (CDS)	4,366
Protein-coding CDS	4,162
Non-protein-coding CDS	204
Genes	4,512

with a coverage of 60x. CheckM v1.2.4 quantified genome contamination and completeness by the *Xanthomonas campestris* CheckM marker set (9). Genome annotation of *Xcc* strain NSTU AG-1 was performed using Prokaryotic Genome Annotation Pipeline (PGAP) v6.10 (10). The genomic features were summarized and represented in Table

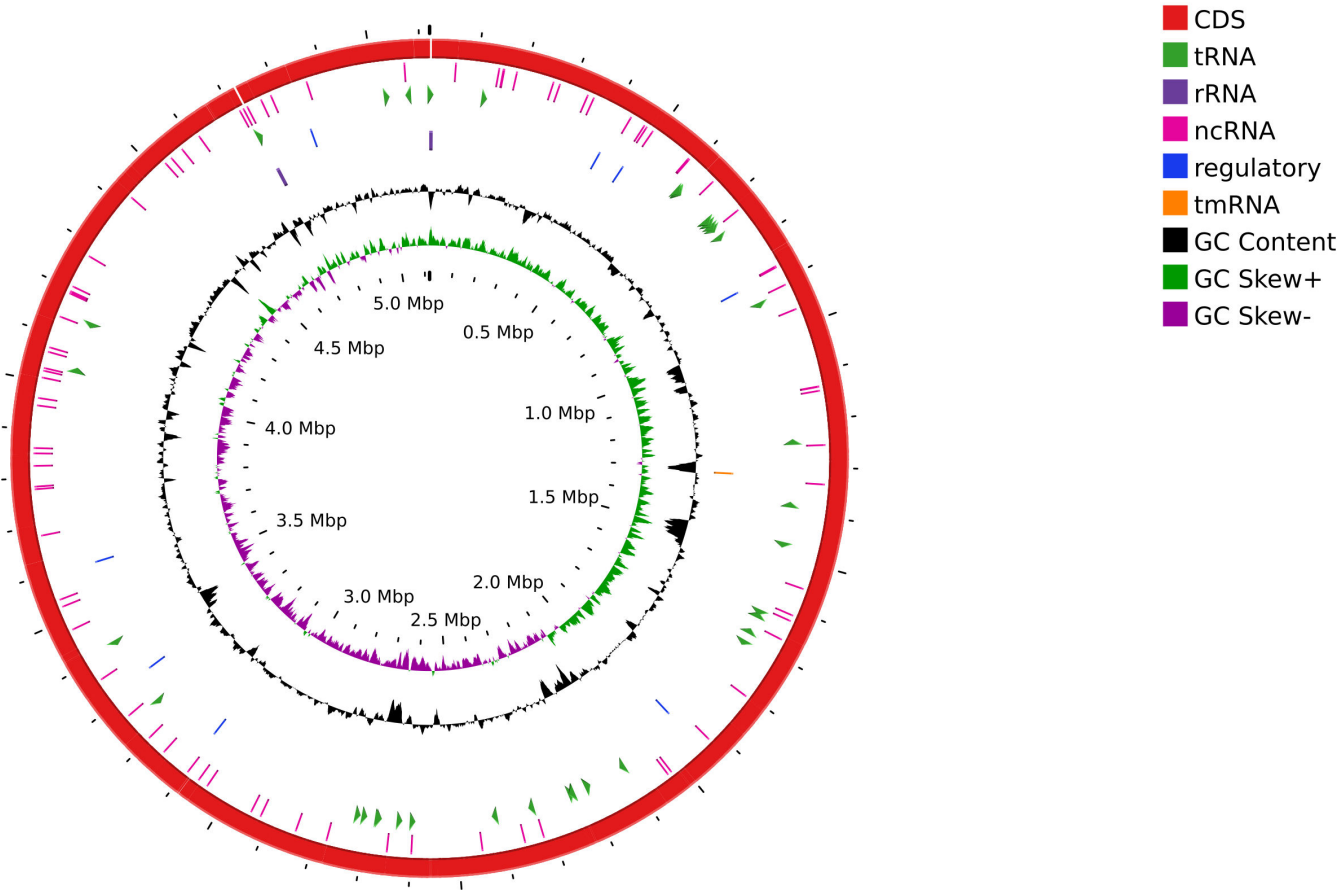


FIG 1 Circular genome map of *Xanthomonas campestris* pv. *campestris* strain NSTU_AG-1 generated using the Proksee server (13). From the outside to the center: circle 1, CDS (red color); circle 2, ncRNA (pink color); circle 3, tRNA (green color); circle 4, regulatory genes (blue color); circle 5, rRNA (violet color); circle 6, tmRNA (deep yellow color); circle 7, GC content (black); and circle 8, GC skew $[(G + C)/(G - C)]$ plotted using a 10-kb window size (purple indicates values below average, while olive represents values above average).

1 and in the circular map (Fig. 1). DNA:DNA hybridization (89.2%; Formula 2: d4) and average nucleotide identity (ANI) (98.75%) between the assembled genome and the closest taxon, *Xcc* strain ATCC 33913 (National Center for Biotechnology Information [NCBI] accession no. [NC_003902.1](https://www.ncbi.nlm.nih.gov/nuclot/NC_003902.1)), using GGDC 3.0 and ANI calculator (<https://www.ezbiocloud.net/tools/ani>), respectively, confirmed it as *Xcc* (11, 12). These genomic data of *Xcc* strain NSTU_AG-1 from Bangladesh will help researchers understand this pathogen and build sustainable disease management techniques.

ACKNOWLEDGMENTS

The Bangladesh Bureau of Educational Information Statistics (BANBEIS), Ministry of Education, Bangladesh is thanked for funding (LS20233103) this research work.

AUTHOR AFFILIATIONS

¹Department of Agriculture, Noakhali Science and Technology University, Noakhali, Bangladesh

²DNA Solution Ltd., Bir Uttam Quazi Nuruzzaman Sarak, Dhaka, Bangladesh

³Department of Microbiology, Noakhali Science and Technology University, Noakhali, Bangladesh

⁴Institute of Biotechnology and Genetic Engineering, Gazipur Agricultural University, Gazipur, Bangladesh

⁵Department of Horticulture, Suncheon National University, Suncheon, Republic of Korea

AUTHOR ORCIDs

Mehede H. Rubel  <http://orcid.org/0000-0001-5872-9361>

FUNDING

Funder	Grant(s)	Author(s)
Bangladesh Bureau of Educational Information and Statistics (BANBEIS)	Project code: LS20233103	Mehede H. Rubel

AUTHOR CONTRIBUTIONS

Arnab Goswami, Formal analysis, Methodology, Validation, Writing – original draft, Writing – review and editing | Maksudur R. Nayem, Data curation, Formal analysis, Software, Visualization | Pijush K. Jhan, Conceptualization, Funding acquisition, Investigation, Project administration, Supervision, Writing – review and editing | Amena Khatun, Investigation, Writing – review and editing | A.B.Z. N. Rahman, Data curation, Formal analysis, Software, Writing – review and editing | Mahmudul Hasan, Software, Writing – review and editing | Swagato Dutta, Visualization, Writing – review and editing | Md T. Islam, Investigation, Visualization, Writing – review and editing | Jong-In Park, Investigation, Writing – review and editing | Mehede H. Rubel, Conceptualization, Funding acquisition, Investigation, Writing – review and editing

DATA AVAILABILITY

The complete genome sequences have been deposited in GenBank under accession number [CP199799.1](https://www.ncbi.nlm.nih.gov/nuclot/CP199799.1) (chromosome). The BioProject accession number is [PRJNA1321718](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1321718), and the BioSample accession number is [SAMN51222686](https://www.ncbi.nlm.nih.gov/biosample/SAMN51222686). Raw reads are available at the NCBI SRA under accession number [SRX30953418](https://www.ncbi.nlm.nih.gov/sra/SRX30953418).

REFERENCES

- Vicente JG, Holub EB. 2013. *Xanthomonas campestris* pv. *campestris* (cause of black rot of crucifers) in the genomic era is still a worldwide threat to brassica crops. *Mol Plant Pathol* 14:2–18. <https://doi.org/10.1111/1/j.1364-3703.2012.00833.x>
- Williams PH. 1981. Black rot: a continuing threat to world crucifers. *Plant Dis* 64:736–742. <https://doi.org/10.1094/PD-64-736>
- Kong C, Chen G, Yang L, Zhuang M, Zhang Y, Wang Y, Ji J, Fang Z, Lv H. 2021. Germplasm screening and inheritance analysis of resistance to

- cabbage black rot in a worldwide collection of cabbage (*Brassica oleracea* var. *capitata*) resources. *Sci Hortic* 288:110234. <https://doi.org/10.1016/j.scienta.2021.110234>
4. Afrin KS, Rahim MA, Park J-I, Natarajan S, Rubel MH, Kim H-T, Nou I-S. 2018. Screening of cabbage (*Brassica oleracea* L.) germplasm for resistance to black rot. *Plant Breed Biotech* 6:30–43. <https://doi.org/10.9787/PBB.2018.6.1.30>
 5. Stanojević N, Živanović B, Milić Komić S, Marković S, Popović Milovanović T, Mitrović P, Sofrenić I, Mutavdžić D, Milinčić D, Pešić M, Jelušić A. 2025. *Bacillus velezensis*: a natural shield against black rot disease on Serbian autochthonous cabbage cv. Futoški. *Chem Biol Technol Agric* 12:133. <https://doi.org/10.1186/s40538-025-00851-6>
 6. Takeuchi K, Mitsuura I. 2020. Complete genome sequences of two strains of *Xanthomonas campestris* pv. *campestris* isolated in Japan. *Microbiol Resour Announc* 9:01239–19. <https://doi.org/10.1128/MRA.01239-19>
 7. 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>
 8. 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>
 9. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>
 10. Tatusova T, DiCuccio M, Badretdin A, Chetvernin 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>
 11. Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 14:1–14. <https://doi.org/10.1186/1471-2105-14-60>
 12. Yoon S-H, Ha S-M, Lim J, Kwon S, Chun J. 2017. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* 110:1281–1286. <https://doi.org/10.1007/s10482-017-0844-4>
 13. Grant JR, Enns E, Marinier E, Mandal A, Herman EK, Chen C-Y, Graham M, Van Domselaar G, Stothard P. 2023. Proksee: in-depth characterization and visualization of bacterial genomes. *Nucleic Acids Res* 51:W484–W492. <https://doi.org/10.1093/nar/gkad326>