

Draft genome sequence of *Cronobacter dublinensis* S21, isolated from a child nutrition formula factory in Bangladesh

Sutapa Bhowmik,^{1,2} Supantha Rivu,³ Md. Latiful Bari,⁴ Sangita Ahmed²

AUTHOR AFFILIATIONS See affiliation list on p. 3.

ABSTRACT Here, we present the draft genome of *Cronobacter dublinensis* S21, isolated from a shoe rack at a child nutrition formula factory in Dhaka, Bangladesh. The genome is 4,564,625 bp with 4,085 predicted coding sequences and G+C content of 57.6%, and displays 40 virulence and 17 antibiotic resistance genes.

KEYWORDS *Cronobacter dublinensis*, draft genome, food processing industry, Bangladesh

Cronobacter dublinensis is a gram-negative bacterium belonging to the genus *Cronobacter* (1). Although no clinical cases of *C. dublinensis* have been reported to date, this bacterium has demonstrated a high capability to invade HEp-2 cells (2). *Cronobacter sakazakii*, another species within the genus, has been reported in several studies in Bangladesh (3–5); however, there is no information regarding *C. dublinensis* from this region. Here, we share the draft genome sequence of *C. dublinensis* S21, which was isolated in 2024 from a shoe rack at a child nutrition formula factory in Dhaka, Bangladesh.

A swab sample was collected from the shoe rack of a child nutrition formula factory in Dhaka (23°42′47.9808″ N, 90°34′36.1704″ E) and transferred immediately to the Department of Microbiology, University of Dhaka. The sample was enriched in Enterobacteria Enrichment Broth (Oxoid, England), followed by streaking on *Enterobacter sakazakii* agar (ESA) (Oxoid) and incubation at 37°C for 48 h (3). A blue colony from ESA was identified as a potential *Cronobacter* sp. microscopically. Later, *gluA* gene-specific polymerase chain reaction presumptively confirmed the isolate as *Cronobacter* sp. Using the Qiagen DNeasy Blood and Tissue Kit (250) (#69504; Qiagen, Hilden, Germany), genomic DNA was extracted from 1.5 mL of an overnight bacterial culture cultivated in Luria broth (Oxoid, Basingstoke, Hampshire, UK) at 37°C at 120 rpm. After ensuring the DNA quality using Nanodrop (ND-1000; Thermo Fisher Scientific, Waltham, MA, USA) and Qubit (Q33238, Thermo Fisher Scientific), it was sent to the International Center for Diarrheal Disease Research, Dhaka, Bangladesh, for whole-genome sequencing utilizing the Illumina NextSeq 2000 platform (Illumina, San Diego, CA, USA). For library preparation, the Illumina DNA Prep Reagent Kit (20060059) (Illumina) was used alongside an automated liquid handler (epMotion 5075) (3).

Upon assessing the quality using FastQC v0.12.1 (6), sequenced fastq files were trimmed using fastp v1.0.1 followed by assembly using SPAdes v4.2.0 (7, 8). Upon discarding contigs less than 500 bp, Prokka v1.14.6 was utilized for annotation (9). Virulence genes were detected using the Virulence Factor Database in ABRicate v1.0.1 on Galaxy with 80% coverage and 70% identity, respectively. Resistance Gene Identifier v6.0.5 in Comprehensive Antibiotic Resistance Database v4.0.1 was applied to detect antibiotic resistance genes (ARGs) (10–13). Later, PlasmidFinder v2.1.6, MobileElementFinder v1.0.3, and CRISPRCasFinder web server were used to detect plasmids, insertion

Editor Silvia T. Cardona, University of Manitoba, Winnipeg, Manitoba, Canada

Address correspondence to Sangita Ahmed, sangita@du.ac.bd.

Sutapa Bhowmik and Supantha Rivu contributed equally to this article. Author order was determined on the basis of seniority.

The authors declare no conflict of interest.

See the funding table on p. 4.

Received 10 January 2026

Accepted 28 March 2026

Published 5 May 2026

Copyright © 2026 Bhowmik 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/).

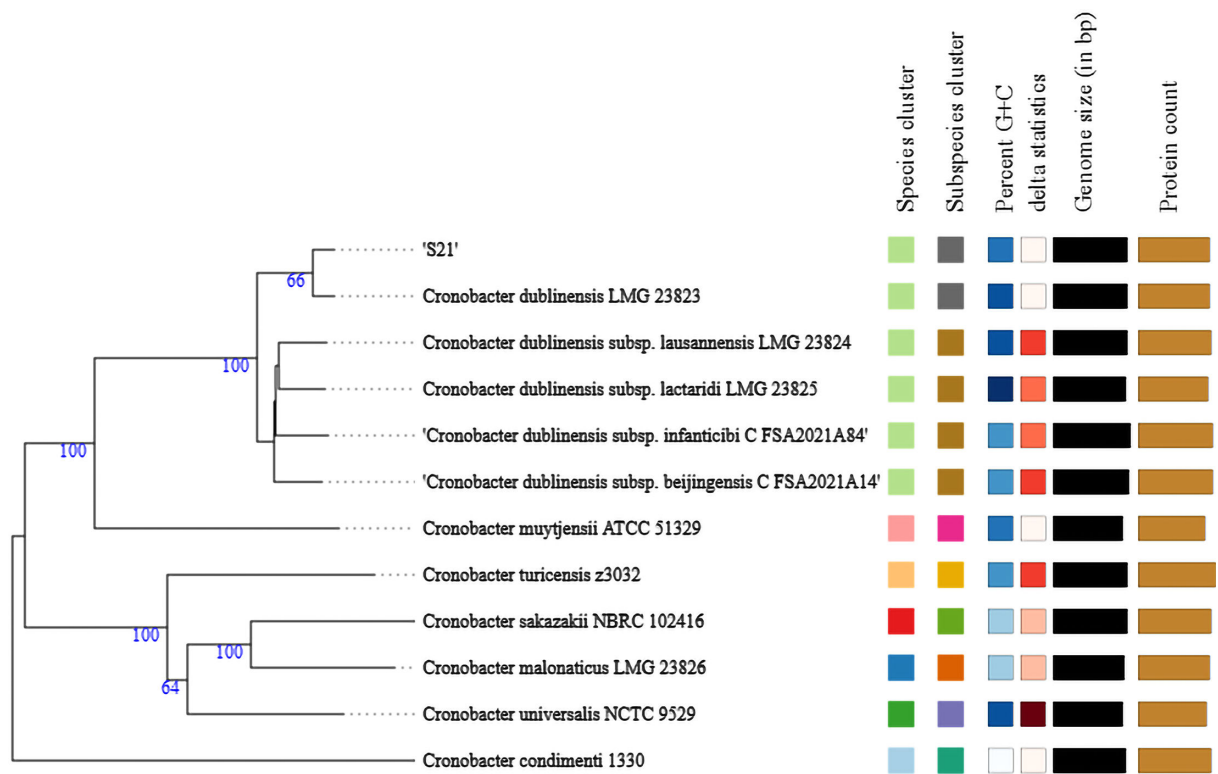


FIG 1 Phylogenetic analysis by Type Strain Genome Server shows that the study genome had the highest similarity with *C. dublinensis*. The whole-genome phylogenetic tree was constructed based on the Genome Blast Distance Phylogeny (GBDP) approach. It inferred with FastME version 2.1.6.1 (17) from GBDP distances calculated from genome sequences. The branch lengths are scaled in terms of GBDP distance formula d5. The numbers above branches are GBDP pseudo-bootstrap support values >60% from 100 replications, with an average branch support of 73.4%. The tree was rooted at the midpoint (18). Here, leaf labels are annotated by affiliation to species clusters, subspecies clusters, percent G+C, delta statistics, genome size (in bp), and protein count (from left to right).

sequences, and CRISPR arrays, respectively (14–16). All these tools were applied using default parameters unless otherwise specified.

The assembly consisted of 57 contigs with a total length of 4,564,625 bp and a G+C content of 57.6% (Table 1). The sequencing depth of the assembled genome was 51.23×. Whole-genome phylogenetic analysis using the Type Strain Genome Server grouped this genome with the *C. dublinensis* LMG 23823 (Fig. 1). Prokka revealed the presence of 4,085 coding DNA sequences, 67 tRNAs, and 3 rRNAs (Table 1). A total of 40 virulence genes, 17 ARGs, 1 plasmid, 2 insertion sequences, and 6 CRISPRs were detected in the genome (Table 1).

TABLE 1 Assembly statistics and important genomic features of *C. dublinensis* S21 from a shoe rack sample in Bangladesh

Parameter	Result for <i>C. dublinensis</i> S21
Sample source	Shoe rack
Sampling location	Child nutrition formula factory, Dhaka, Bangladesh
Sample type	Cotton swab
Year of isolation	2024
No. of reads	855,514
Total output (Mbp)	126.5
Type of reads	Paired end
Read length (bp)	35–151
Genome length (bp)	4,564,625
No. of contigs	57
GC content (%)	57.6
N50 (bp)	285,910
L50	7
Sequencing depth (×)	51.23
Coding DNA sequences	4,085
tRNAs	67
tmRNA	1
rRNAs	3
Antibiotic resistance genes	17 (<i>acrA</i> , <i>adeF</i> , <i>CRP</i> , <i>emrR</i> , <i>FosA8</i> , <i>H-NS</i> , <i>KpnE</i> , <i>KpnF</i> , <i>KpnH</i> , <i>marA</i> , <i>msbA</i> , <i>PBP3</i> , <i>qacG</i> , <i>rsmA</i> , <i>vanG</i> , <i>EF-Tu</i> mutants, <i>AcrAB-TolC</i> with <i>MarR</i> mutations)
Virulence genes	40 (<i>cheB</i> , <i>cheD</i> , <i>cheW</i> , <i>cheY</i> , <i>cheZ</i> , <i>csgD</i> , <i>csgF</i> , <i>csgG</i> , <i>entA</i> , <i>entB</i> , <i>entE</i> , <i>entS</i> , <i>fepA</i> , <i>fepB</i> , <i>fepC</i> , <i>fepD</i> , <i>fepG</i> , <i>flgB</i> , <i>flgC</i> , <i>flgG</i> , <i>flgH</i> , <i>flgI</i> , <i>flhA</i> , <i>flhC</i> , <i>fliA</i> , <i>fliG</i> , <i>fliI</i> , <i>fliM</i> , <i>fliP</i> , <i>fliS</i> , <i>hsiB1/vipA</i> , <i>hsiC1/vipB</i> , <i>htpB</i> , <i>llpA</i> , <i>kdsA</i> , <i>kpsD</i> , <i>luxS</i> , <i>mgtB</i> , <i>motA</i> , <i>ompA</i>)
Plasmid	1 [Col(pHAD28) (KU674895)]
Insertion sequences	2 (ISEhe3, ISRaql1)
CRISPR array	6
Accession no.	NZ_JBSXVU0000000000

ACKNOWLEDGMENTS

This work was supported by the University Grants Commission, Bangladesh (Grant no. Biology-34/2023-24).

AUTHOR AFFILIATIONS

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

²Department of Microbiology, University of Dhaka, Dhaka, Bangladesh

³Department of Microbiology, Notre Dame University Bangladesh, Dhaka, Bangladesh

⁴Food, Nutrition and Agriculture Research Division, Centre for Advanced Research in Sciences, University of Dhaka, Dhaka, Bangladesh

AUTHOR ORCIDs

Supantha Rivu  <http://orcid.org/0009-0000-9713-0183>

Sangita Ahmed  <http://orcid.org/0000-0002-0774-4542>

FUNDING

Funder	Grant(s)	Author(s)
University Grant Commission, Bangladesh	Biology-34/2023-24	Md. Latiful Bari
University Grants Commission, Bangladesh	Biology-34/2023-24	Sangita Ahmed

AUTHOR CONTRIBUTIONS

Sutapa Bhowmik, Data curation, Formal analysis, Methodology, Software, Writing – original draft | Supantha Rivu, Formal analysis, Methodology, Software, Writing – original draft | Md. Latiful Bari, Conceptualization, Funding acquisition, Supervision | Sangita Ahmed, Conceptualization, Funding acquisition, Supervision, Validation, Writing – review and editing

DATA AVAILABILITY

The whole-genome sequencing data for *C. dublinensis* S21 have been submitted to GenBank with accession no. [NZ_LBSXVU000000000](https://www.ncbi.nlm.nih.gov/nuclseq/NZ_LBSXVU000000000) under BioProject no. [PRJNA1380941](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1380941) (BioSample accession no. [SAMN53932570](https://www.ncbi.nlm.nih.gov/biosample/SAMN53932570) and SRA accession no. [SRR37066346](https://www.ncbi.nlm.nih.gov/sra/SRR37066346)).

REFERENCES

- Kucerova E, Joseph S, Forsythe S. 2011. The *Cronobacter* genus: ubiquity and diversity. *Qual. Assur. Saf. Crop. Foods* 3:104–122. <https://doi.org/10.1111/j.1757-837X.2011.00104.x>
- Cruz A, Xicohtencatl-Cortes J, González-Pedrajo B, Bobadilla M, Eslava C, Rosas I. 2011. Virulence traits in *Cronobacter* species isolated from different sources. *Can J Microbiol* 57:735–744. <https://doi.org/10.1139/w11-063>
- Bhowmik S, Rivu S, Bari ML, Ahmed S. 2025. Genome mining of *Cronobacter sakazakii* in Bangladesh reveals the occurrence of high-risk ST83 and rare ST789 lineages. *Pathogens* 14:1220. <https://doi.org/10.3390/pathogens14121220>
- Fakruddin M, Rahaman M, Ahmed MM, Hoque MM. 2014. Stress tolerant virulent strains of *Cronobacter sakazakii* from food. *Biol Res* 47:63. <https://doi.org/10.1186/0717-6287-47-63>
- Hoque A, Ahmed T, Shahidullah M, Hossain A, Mannan A, Noor K, Nahar K, Ilias M, Ahmed D. 2010. Isolation and molecular identification of *Cronobacter* spp. from powdered infant formula (PIF) in Bangladesh. *Int J Food Microbiol* 142:375–378. <https://doi.org/10.1016/j.ijfoodmicro.2010.07.019>
- Andrews S. 2025. FastQC a quality control tool for high throughput sequence data [Internet]. Available from: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. 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>
- Chen S, Zhou Y, Chen Y, Gu J. 2018. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>
- Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>
- Liu B, Zheng D, Jin Q, Chen L, Yang J. 2019. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Res* 47:D687–D692. <https://doi.org/10.1093/nar/gky1080>
- McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ, et al. 2013. The comprehensive antibiotic resistance database. *Antimicrob Agents Chemother* 57:3348–3357. <https://doi.org/10.1128/AAC.00419-13>
- Seemann T. 2016. ABRicate: mass screening of contigs for antibiotic resistance genes [Internet]. Available from: <https://github.com/tseeman/abricate>
- The Galaxy Community. 2024. The galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update. *Nucleic Acids Res* 52:W83–W94. <https://doi.org/10.1093/nar/gkaf410>
- Johansson MHK, Bortolaia V, Tansirichaiya S, Aarestrup FM, Roberts AP, Petersen TN. 2021. Detection of mobile genetic elements associated with antibiotic resistance in *Salmonella* enterica using a newly developed web tool: MobileElementFinder. *J Antimicrob Chemother* 76:101–109. <https://doi.org/10.1093/jac/dkaa390>
- Couvin D, Bernheim A, Toffano-Nioche C, Touchon M, Michalik J, Néron B, et al. 2018. CRISPRCasFinder, an update of CRISPRFinder, includes a portable version, enhanced performance and integrates search for cas proteins. *Nucleic Acids Res* 46:W246–W251. <https://doi.org/10.1093/nar/gky425>
- Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, Villa L, et al. 2014. *In silico* detection and typing of plasmids using plasmid-finder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother* 58:3895–3903. <https://doi.org/10.1128/AAC.02412-14>
- Lefort V, Desper R, Gascuel O. 2015. FastME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program. *Mol Biol Evol* 32:2798–2800. <https://doi.org/10.1093/molbev/msv150>
- Farris JS. 1972. Estimating phylogenetic trees from distance matrices. *Am Nat* 106:645–668. <https://doi.org/10.1086/282802>