Data in Brief 57 (2024) 111133

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Draft genome sequence data of *Serratia marcescens* strain harboring bla_{NDM-7} from Dhaka, Bangladesh

K.M. Shayerul Abedin Shayer[a], Sabbir R. Shuvo[a], Ishrat Jabeen[a], Mahmud Hossain[b], Sohidul Islam[a,*]

[a] Department of Biochemistry & Microbiology, North South University, Dhaka, Bangladesh
[b] Department of Biochemistry & Molecular Biology, University of Dhaka, Dhaka, Bangladesh

ARTICLE INFO

Article history:
Received 25 October 2024
Revised 6 November 2024
Accepted 7 November 2024
Available online 14 November 2024

Dataset link: Pathogen: clinical or host-associated sample from Serratia marcescens (Original data)

Keywords:
Serratia marcescens
WGS
Multidrug resistance
NDM-7
Phylogenomic analysis

ABSTRACT

Here, the draft genome sequence of a multi-drug resistant (MDR) *Serratia marcescens* strain BMD28, isolated from a clinical source from Dhaka, Bangladesh, has been reported. The sequence raw read files were generated using Illumina sequencing technology utilizing genomic DNA from the pure culture of this strain. The strain has a genome size of around 5.4 million base pairs, a GC content of 59.70 %, and 5,141 coding sequences. We conducted genomic studies using several bioinformatics tools focusing on resistance genes, virulence factors, toxin-antitoxin systems, and pangenome analysis. Strain BMD28 harbored the bla_{NDM-7} gene in an IncX3 plasmid. A phylogenomic study with *S. marcescens* strains isolated worldwide revealed that our strain is in the same clade as other strains reported in Bangladesh. The data can be used primarily to understand the genomic content, epidemiology, and evolution of *S. marcescens* in Bangladesh. The genome sequence data of BMD28 has been deposited in the NCBI database under BioSample accession number SAMN41260295.

* Corresponding author.
E-mail address: sohidul.islam@northsouth.edu (S. Islam).

https://doi.org/10.1016/j.dib.2024.111133

Specifications Table

Subject	Biological Sciences
Specific subject area	Omics: Genomics
Type of data	Raw, table, figure, analyzed, deposited
Data collection	The bacterial strain was isolated from a urine sample of a UTI patient in Dhaka, Bangladesh. Genomic DNA was extracted, and paired-end raw reads were generated using the Illumina™ NextSeq500 platform. Genome assembly and annotation were done on the BV-BRC server, with a genomic map constructed using Proksee. The isolate's MLST profile was identified through PubMLST. Contigs were analyzed for antibiotic resistance genes, virulence factors, and toxin-antitoxin systems using CARD, VFDB, and TASmania, respectively. Orthologous protein clusters were compared via Orthovenn3. Phylogenomic trees were generated using TYGS and CSIPhylogeny based on WGS and SNP and visualized in iTOL, while BPGA was used for Pan-genome analysis.
Data source location	Institution • Dhaka Central International Medical College Hospital City/Town/Region: DhakaCountry: BangladeshLatitude: 23.8041° N, Longitude: 90.4152° E
Data accessibility	Repository name: NCBI BioSample Data identification number: SAMN41260295 Direct URL to data: https://www.ncbi.nlm.nih.gov/biosample/SAMN41260295
Related research article	None

1. Value of the data

- This study presents the draft genome sequence of a non-pigmented, multidrug-resistant *S. marcescens* strain, BMD28 (SM BMD28), which provides valuable insights into genomic mechanisms of multidrug resistance (MDR) and the species' epidemiology.
- The genome data of SM BMD28 includes one of the first reports of the bla_{NDM-7} gene in a Bangladeshi *S. marcescens* strain, significantly contributing to the understanding of the dissemination and evolution of carbapenem-resistant Enterobacteriaceae.
- The draft genome sequence data may serve as a resource for comparative studies to elucidate patterns of multidrug resistance, virulence factors, phylogenomic relationships, and pangenome characteristics in *S. marcescens* in Bangladesh.

2. Background

S. marcescens has emerged as a notable nosocomial pathogen, contributing significantly to the incidence of healthcare-associated infections [1]. The prevalence of *S. marcescens* in hospital environments and clinical samples in Bangladesh, as well as in other low- and middle-income countries, has been increasing [2,3]. Therefore, exploring the genomic characteristics of MDR *S. marcescens* through whole genome sequencing (WGS) could provide valuable insights into the potential factors that may influence the emergence and spread of MDR *S. marcescens* in Bangladesh.

This study aims to characterize the genomic content of the MDR strain SM BMD28 isolated in Bangladesh alongside a comparative analysis of its antimicrobial resistance traits, virulence factors, and phylogenomic relationships.

3. Data Description

3.1. Antibiotic susceptibility profile of SM BMD28

The strain was identified as non-pigmented (Fig. S1); and MDR by conferring resistance to the majority of tested antibiotics, with the exception of those in the fluoroquinolone class (Table 1).

3.2. Genomic features of SM BMD28

The SM BMD28 strain has a GC content of 59.70 % and a genome size of approximately 5.4 million base pairs, comprising 5141 coding sequences (Table 2). The local alignment of the Bangladeshi *S. marcescens* strains against the reference genome of *S. marcescens* ELP1.10 is visualized in Fig. 1. The strains *S. marcescens* BK-E759 and BC-E65 were used for this alignment to represent at least one cluster in the average nucleotide identity (ANI) matrix (Fig. S2).

Table 1
Antibiotic sensitivity pattern of the Bangladeshi *S. marcescens* isolate SM BMD28.

Antibiotic class	Antibiotic name	Interpretation
Aminoglycosides	AMK	R
	GEN	R
Carbapenem	IPM	R
	MEM	R
Cephalosporins	CTX	R
	FEP	R
Fluroquinolones	CIP	S
	LEV	S
Sulfonamide	SXT	R
Penicillin + beta-lactamase inhibitors	TZP	I
Polymyxin[a]	PMB	R
	COL	R

R=Resistant, *S*=Sensitive, *I*=Intermediate
[a] Determined by broth microdilution (BMD) method generating Minimum Inhibitory Concentration (MIC) value for PMB (32 mg/L) and COL (128 mg/L) according to CLSI guidelines.

Table 2
Genomic characteristics of SM BMD28.

Attribute	SM BMD 28
Place of Isolation	Dhaka, Bangladesh
Isolation Source	Urine
Genome Size (Mbp)	5.4
Number of Contigs	88
[a]N50	270,090
[b]L50	8
GC (%)	59.70
Coding Genes	5141
tRNA	79
rRNA	5
MLST	ST366
Bio-sample ID	SAMN41260295

** [a]N50 = Half of the genome assembly is contained in contigs equal to or larger than this value.
[b] L50 = smallest number of contigs (each with its length) in the genome assembly needed to cover approximately half of the total genome size.

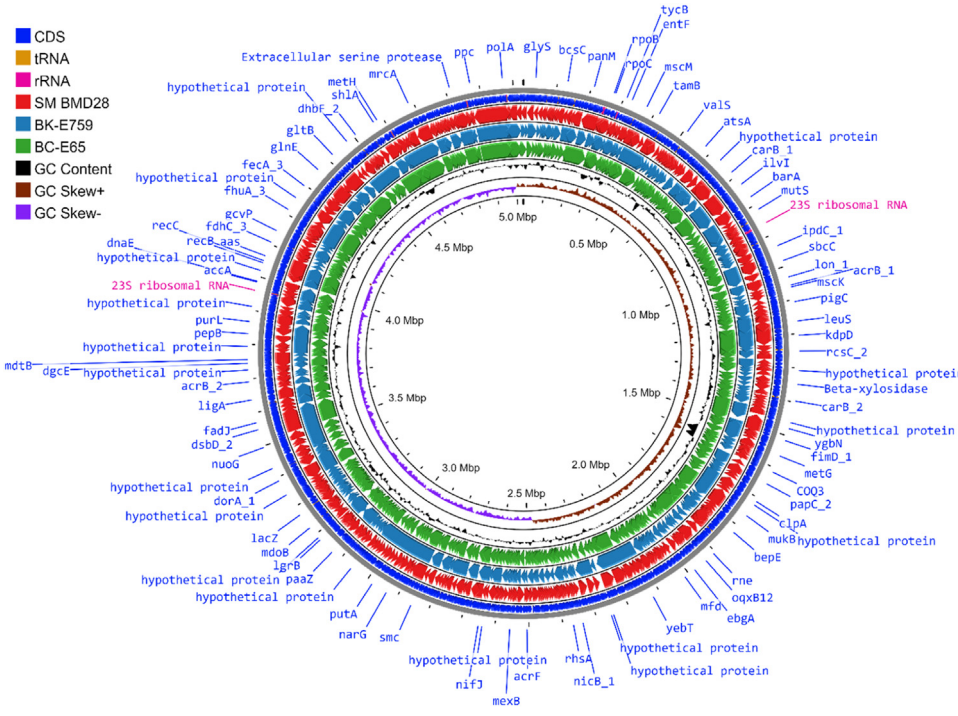

Fig. 1. Sequence alignment of the isolated SM BMD28 using *S. marcescens* strain ELP1.10 as the reference genome. The gaps in each circular genome represent the missing regions identified in BLAST analysis. The inner circle represents the sequence clockwise. The blue arrows denote the annotated coding DNA sequences (CDS). The brown peaks correspond to the positive GC skew, while the purple peaks represent the negative GC skew. Black peaks indicate the GC content. (For interpretation of the color references in this figure legend, the reader is referred to the web version of this article.)

3.3. *Prediction of antibiotic resistance and virulence genes in* S. marcescens *strains*

Utilizing the Comprehensive Antibiotic Resistance Database (CARD), we identified antibiotic resistance genes present in the studied strains and conducted a comparative analysis with the reference strain ELP1.10, as well as representative strains from each Bangladeshi *S. marcescens* cluster determined by ANI (Fig. 2A). Our findings indicate that SM BMD28 possesses genes such as $bla_{NDM\text{-}7}$ and $bla_{SHV\text{-}134}$. Additionally, using the Virulence Factor Database (VFDB), the virulence profile of SM BMD28 shows substantial similarity to that of the reference strain (Fig. 2B).

3.4. *IncX3 plasmid harbored the* $bla_{NDM\text{-}7}$ *gene*

The $bla_{NDM\text{-}7}$ was found in contig number 26, which is also predicted by the CARD database (Fig. 2A). Contig 26 was later identified as IncX3 plasmid by PlasmidFinder 2.0 (https://cge.food.dtu.dk/services/PlasmidFinder-2.0/). This IncX3 plasmid was aligned with complete *S. marcescens* plasmid p_dmsm540_NDM7 (CP095678.1) (Fig. 3). In the ClustalW(https://www.ebi.ac.uk/jdispatcher/msa/clustalo) analysis comparing a $bla_{NDM\text{-}1}$ reference sequence (NG_049326.1) with both the *S. marcescens* $bla_{NDM\text{-}7}$ sequences from p_dmsm540_NDM7 and BMD28 IncX3 plasmid, only two nucleotide differences were observed at positions 488 and 560. In the $bla_{NDM\text{-}1}$ sequence, these positions were occupied by G and A, respectively, whereas in the $bla_{NDM\text{-}7}$ sequences, they were replaced by A and C (Fig. S3).

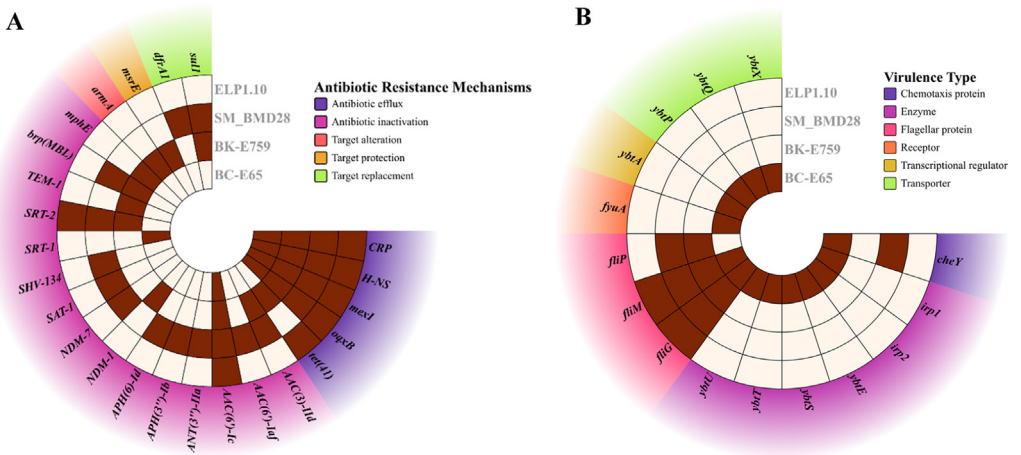

Fig. 2. Antibiotic resistance genes (A) and virulence factors (B) of SM BMD28 along with two other Bangladeshi *S. marcescens* strains from each ANI-matched cluster and the reference strain ELP1.10. The brown color represents the presence of the genes, and the white color represents the absence of the genes. (For interpretation of the color references in this figure legend, the reader is referred to the web version of this article.)

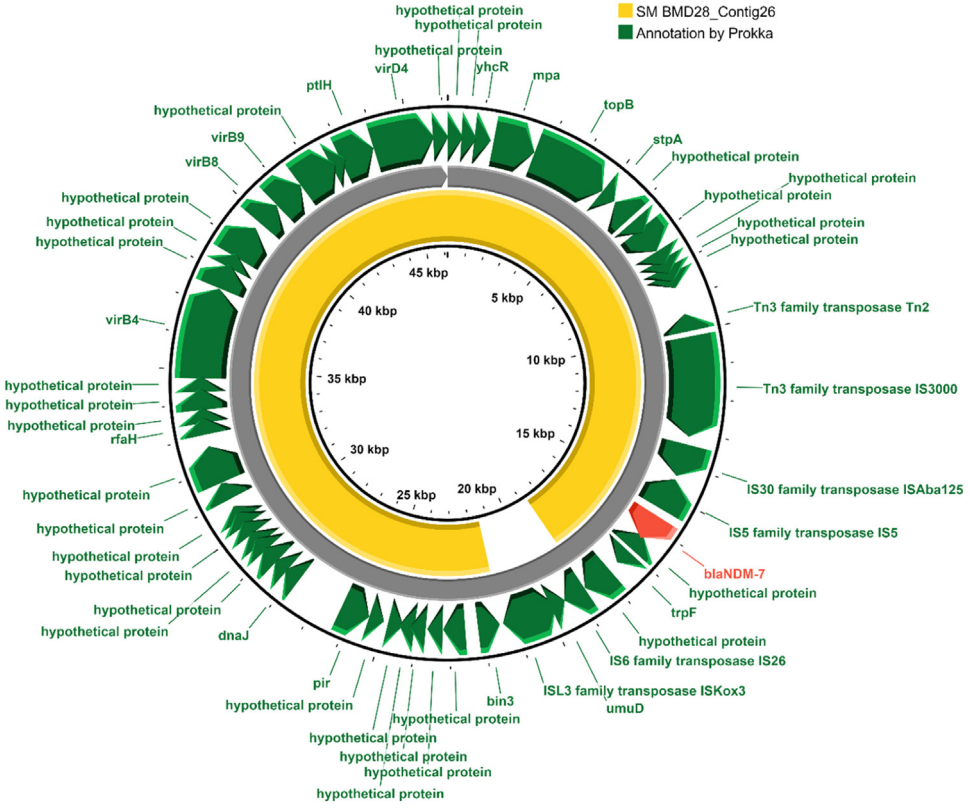

Fig. 3. Sequence alignment of the IncX3 plasmid from BMD28 using *S. marcescens* complete plasmid p_dmsm540_NDM7 (CP095678.1) as a template.

3.5. Prediction of toxin-antitoxin (TA) systems

The TA systems were examined using the TASer search tool, which is integrated with the TASmania database and illustrated as a heatmap in Fig. 4 [4]. The type II toxin-antitoxin (TA) system represents the predominant complete TA systems identified in Bangladeshi *S. marcescens* strains, including ParDE, RelBE. Among the identified TA system genes, we predicted the presence of numerous type II, type IV, and type I toxin genes, as well as type II antitoxin genes, across the studied strains. A notable difference in the composition of TA system genes was observed between SM BMD28 and the reference strain ELP1.10 (Fig. 4).

3.6. Orthologous protein cluster analysis

The SM BMD28, alongside two Bangladeshi strains were analyzed in comparison with the reference strain ELP1.10 to identify unique gene families (Fig. 5). Across all strains, a total of 3877 common orthologous clusters were identified. The strain SM BMD28 exhibited 38 unique gene clusters that were absent from BK-E759, BC-E65, and ELP1.10.

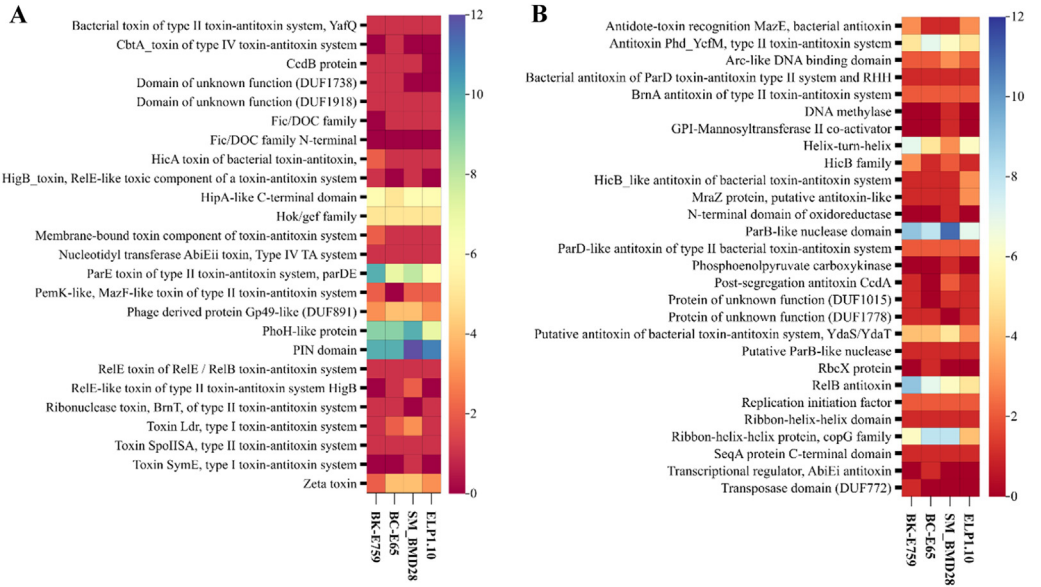

Fig. 4. Toxin-Antitoxin system of the *S. marcescens* strains and the reference strain ELP1.10. Strain names are mentioned on the X-axis. The predicted toxin and antitoxin genes can be seen in the Y axis. Varying colors represent different numbers of toxin (A) and antitoxin (B) genes. (For interpretation of the color references in this figure legend, the reader is referred to the web version of this article.)

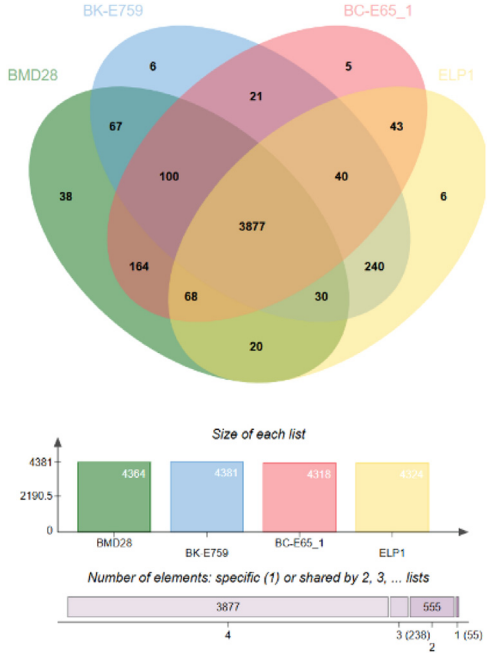

Fig. 5. *S. marcescens* strains are depicted in a Venn diagram with distinct and common orthogroups/gene families.

3.7. Phylogenomic analysis

A phylogenomic tree was constructed to compare the genomic similarities among 50 *S. marcescens* strains worldwide (including the Bangladeshi isolates) selected based on the ANI matrix. (Fig. 6A). The strain SM BMD28 is present in a separate node from other Bangladeshi strains. A SNP cladogram of global *S. marcescens* strains, including the studied strain, was also constructed (Fig. 6B).

3.8. Pan-genome analysis

Pangenome analysis of 13 ANI-matched Bangladeshi *S. marcescens* strains identified 3779 core genes shared among all strains. Notably, strains BMD28, IUBTGEM_3, and RMCH-M16-N harbored the highest number of unique genes, with 575, 172, and 100 unique genes, respectively. These three strains also exhibited the highest number of missing genes relative to the other Bangladeshi strains (Table S2). The core–pan plot analysis shows that the *S. marcescens* pan-genome remains "open." However, it is approaching closure as more genomes are incorporated, having minimal impact on the overall pan-genome size (Fig. 7). In the power-law equation $f(x) = a.x^b$, the *B*pan value, representing the total expansion rate of gene families, is determined to be 0.10330 (i.e., <1), associated with the stabilization of the pan-genome [5].

4. Experimental Design, Materials and Method

Isolation and phenotypic characterization of SM BMD28

The *S. marcescens* isolate was collected from urine samples was isolated on Nutrient Agar (Condalab, Spain), followed by subculture on both MacConkey Agar (Condalab, Spain), and Nu-

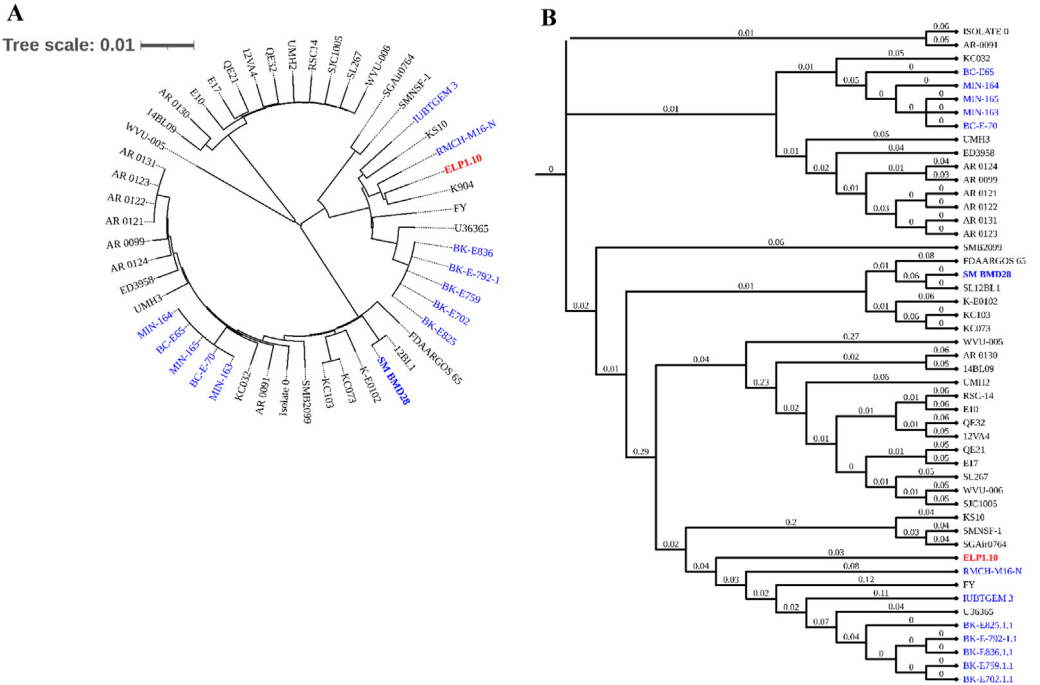

Fig. 6. Phylogenomic tree based on (A) WGS (B) SNP of ANI matched 50 *S. marcescens* strains. The reference strain is labeled red; all the Bangladeshi strains, including SM BMD28, are labeled blue. The corresponding GenBank ID of all these strains is mentioned in Table S1. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

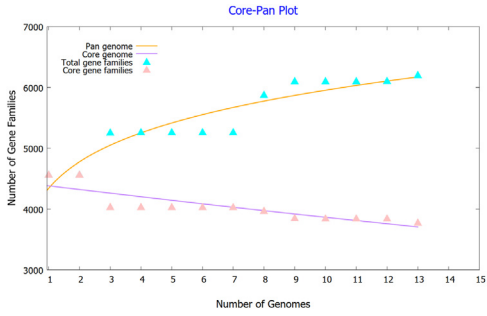

Fig. 7. Pan-genome analysis default core–pan plot of the 13 ANI-matched *S. marcescens* strains.

trient Agar (Condalab, Spain) to observe their distinct characteristics. The isolate was later confirmed as *S. marcescens* by 16 s rDNA sequencing. The confirmed strain then further tested to identify their antimicrobial resistance pattern by using the Kirby-Bauer disk diffusion technique [6] using amikacin (AMK 30 µg), gentamicin (GEN 10 µg), meropenem (MEM 10 µg), imipenem (IMP 10 µg), ceftriaxone (CRO 30 µg), cefepime (FEP 5 µg), ciprofloxacin (CIP 5 µg), levofloxacin (LEV 5 µg), trimethoprim/Sulfamethoxazole (SXT 25 µg), tazobactam with piperacillin (TZP 110 µg) (Bioanalyse, Turkey). The minimum inhibitory concentration (MIC) of all *S. marcescens* strains against colistin and polymyxin-B (Sigma-Aldrich, USA) was calculated quantitatively using broth-microdilution (BMD) in accordance with CLSI recommendations [7].

4.1. Genome assembly and annotation

A single bacterial colony was cultured in BD Difco™ (Becton, Dickinson and Company, USA) LB (Luria-Bertani) broth and incubated overnight at 37 °C in a Memmert IN160 natural convection incubator (Memmert GmbH, Germany). The density of the broth culture was measured using a BioSan DEN-1 densitometer (BioSan, Latvia). DNA was subsequently extracted from the culture using the Qiagen DNeasy Blood & Tissue Kit (Qiagen, USA), following the manufacturer's instructions. The purity of the extracted DNA was checked using a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific, USA), while the DNA concentration was measured with a Qubit 4.0 fluorometer (Thermo Fisher Scientific, USA). Library preparation was performed using the Illumina DNA Prep Reagent Kit (Illumina, USA) in accordance with the manufacturer's protocol. Raw sequencing reads were generated on a NextSeq 500 platform (Illumina™, USA) at icddr, b Genome Centre, Dhaka, Bangladesh. Genome assembly and quality control were carried out using Unicycler 0.4.8.0 [8], Quast 5.0.2 [9], Samtools 1.17 [10], and BUSCO 5.7.1 [11]. Annotations were conducted using the Rapid Annotation using Subsystem Technology toolkit (RASTtk) [12], The assembly and annotation processes were supported by the Bacterial and Viral Bioinformatics Resource Center (BV-BRC) [13]. A circular genome map representing the local alignment of the strains was generated using Proksee [14].

4.2. Genomic characterization, prediction of antibiotic resistance genes, virulence, and toxin-antitoxin systems

Multilocus sequence typing (MLST) of the SM BMD28 isolate was performed by identifying sequence variants in seven housekeeping genes using the PubMLST web server (https://pubmlst.org/). The assembled draft genome was then analyzed using Abricate version 1.0.0 [15] to identify antibiotic resistance genes and virulence factors. For this analysis, the Comprehensive Antibiotic Resistance Database (CARD) [16] was used to screen for resistance genes,

while the Virulence Factor Database (VFDB) [17] was employed to detect virulence factors. Toxin-antitoxin systems presence was predicted using Toxin-Antitoxin Systems Mania (TASmania) [4]. All heatmaps were generated using ChiPlot (https://www.chiplot.online/). All analyses were carried out in the respective webservers/tools using the default parameters.

4.3. Comparative analysis

Genome selection and phylogenomic tree generation were performed according to the method described by Tabassum et al. [18]. Average nucleotide identity (ANI) analysis was conducted using FastANI version 1.34 [19]. Orthologous protein cluster comparison was performed between the reference strain ELP1.10 (Accession: NZ_CP127881.1) and representative strains from each Bangladeshi *S. marcescens* cluster, as determined by ANI. These comparisons were performed using the OrthoVenn3 web interface, employing the OrthoFinder algorithm with an E-value threshold of 1e-5 and an inflation parameter of 1.5 [20]. Phylogenomic analyses were carried out in the Type Strain Genome Server (TYGS) for a whole genome-based taxonomic analysis [21]. Subsequently, a phylogenomic assessment based on whole genome single nucleotide polymorphisms (SNPs) was performed using CSIPhylogeny version 1.4, with ELP1.10 as the reference genome [22]. The resulting phylogenomic trees were visualized using the Interactive Tree of Life (iTOL) web interface [https://itol.embl.de/]. The pan-genome analysis was conducted using the Bacterial Pan Genome Analysis (BPGA) v1.3 tool [23]. All analyses were carried out in the respective webservers/tools using the default parameters unless mentioned otherwise.

Limitations

Not Applicable.

Ethics Statement

All authors have read and adhered to the ethical requirements for Data in Brief and confirm that the current work does not involve human subjects, animal experiments, or any data collected from social media platforms. All procedures were approved by the North South University Institutional Review Board/Ethics Review Committee (IRB Approval Number: 2019/OR-NSU/IRB-No.0504).

CRediT Author Statement

K M Shayerul Abedin Shayer: Writing – original draft, Visualization; **Sabbir R. Shuvo:** Writing – review & editing; **Ishrat Jabeen:** Writing – review & editing; **Mahmud Hossain**: Writing - review & editing**; Sohidul Islam:** Conceptualization, Methodology, Supervision, Writing – review & editing, Funding acquisition.

Data Availability

Pathogen: clinical or host-associated sample from Serratia marcescens (Original data) (NCBI BioSample).

Acknowledgments

This work has been supported by a grant from North South University CTRGC (NSU-RP-18-057) to SI. We sincerely thank Mr. Md. Ismail Hossain for his assistance with strain collection, Ms. Sadia Ferdous for her work on the initial studies involving the strain, and Ms. Afia Anjum for her careful proofreading of the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary Materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dib.2024.111133.

References

[1] I. Cosimato, B. Santella, S. Rufolo, P. Sabatini, M. Galdiero, M. Capunzo, G. Boccia, V. Folliero, G. Franci, Current epidemiological status and antibiotic resistance profile of *Serratia marcescens*, Antibiotics 13 (2024) 323, doi:10.3390/antibiotics13040323.

[2] M. Nieto-Rosado, K. Sands, E.A.R. Portal, K.M. Thomson, M.J. Carvalho, J. Mathias, R. Milton, C. Dyer, C. Akpulu, I. Boostrom, P. Hogan, H. Saif, A.D. Sanches Ferreira, T. Hender, B. Portal, R. Andrews, W.J. Watkins, R. Zahra, H. Shirazi, A. Muhammad, S.N. Ullah, M.H. Jan, S. Akif, K.C. Iregbu, F. Modibbo, S. Uwaezuoke, L. Audu, C.P. Edwin, A.H. Yusuf, A. Adeleye, A.S. Mukkadas, J.B. Mazarati, A. Rucogoza, L. Gaju, S. Mehtar, A.N.H. Bulabula, A. Whitelaw, L. Roberts, G. Chan, D. Bekele, S. Solomon, M. Abayneh, G. Metaferia, T.R. Walsh, Colonisation of hospital surfaces from low- and middle-income countries by extended spectrum β-lactamase- and carbapenemase-producing bacteria, Nat. Commun. 15 (2024) 2758, doi:10.1038/s41467-024-46684-z.

[3] K. Sands, M.J. Carvalho, E. Portal, K. Thomson, C. Dyer, C. Akpulu, R. Andrews, A. Ferreira, D. Gillespie, T. Hender, K. Hood, J. Mathias, R. Milton, M. Nieto, K. Taiyari, G.J. Chan, D. Bekele, S. Solomon, S. Basu, P. Chattopadhyay, S. Mukherjee, K. Iregbu, F. Modibbo, S. Uwaezuoke, R. Zahra, H. Shirazi, A. Muhammad, J.-B. Mazarati, A. Rucogoza, L. Gaju, S. Mehtar, A.N.H. Bulabula, A. Whitelaw, T.R. Walsh, Characterization of antimicrobial-resistant Gram-negative bacteria that cause neonatal sepsis in seven low- and middle-income countries, Nat. Microbiol. 6 (2021) 512–523, doi:10.1038/s41564-021-00870-7.

[4] H. Akarsu, P. Bordes, M. Mansour, D.-J. Bigot, P. Genevaux, L. Falquet, TASmania: a bacterial Toxin-Antitoxin Systems database, PLoS Comput. Biol. 15 (2019) e1006946, doi:10.1371/journal.pcbi.1006946.

[5] A. Anjum, J. Tabassum, S. Islam, A.K.M.I. Hassan, I. Jabeen, S.R. Shuvo, Deciphering the genomic character of the multidrug-resistant *Staphylococcus aureus* from Dhaka, Bangladesh, AIMS Microbiol. 10 (2024) 833–858, doi:10.3934/microbiol.2024036.

[6] J. Hudzicki, in: Kirby-Bauer Disk Diffusion Susceptibility Test Protocol, American Society For Microbiology, 2012, pp. 1–13. https://www.asm.org/Protocols/Kirby-Bauer-Disk-Diffusion-Susceptibility-Test-Pro.

[7] CLSIPerformance Standards for Antimicrobial Susceptibility Testing, 31st ed., Clinical and Laboratory Standards Institute, 2019, doi:10.1007/978-3-662-48986-4_300418.

[8] R.R. Wick, L.M. Judd, C.L. Gorrie, K.E. Holt, Unicycler: resolving bacterial genome assemblies from short and long sequencing reads, PLoS Comput. Biol. 13 (2017) e1005595, doi:10.1371/journal.pcbi.1005595.

[9] A. Gurevich, V. Saveliev, N. Vyahhi, G. Tesler, QUAST: quality assessment tool for genome assemblies, Bioinformatics 29 (2013) 1072–1075, doi:10.1093/bioinformatics/btt086.

[10] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, The sequence alignment/map format and SAMtools, Bioinformatics 25 (2009) 2078–2079, doi:10.1093/bioinformatics/btp352.

[11] F.A. Simão, R.M. Waterhouse, P. Ioannidis, E.V. Kriventseva, E.M. Zdobnov, BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs, Bioinformatics 31 (19) (2015) 3210–3212, doi:10.1093/bioinformatics/btv351.

[12] T. Brettin, J.J. Davis, T. Disz, R.A. Edwards, S. Gerdes, G.J. Olsen, R. Olson, R. Overbeek, B. Parrello, G.D. Pusch, M. Shukla, J.A. Thomason, R. Stevens, V. Vonstein, A.R. Wattam, F. Xia, RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes, Sci. Rep. 5 (2015) 8365, doi:10.1038/srep08365.

[13] R.D. Olson, R. Assaf, T. Brettin, N. Conrad, C. Cucinell, J.J. Davis, D.M. Dempsey, A. Dickerman, E.M. Dietrich, R.W. Kenyon, M. Kuscuoglu, E.J. Lefkowitz, J. Lu, D. Machi, C. Macken, C. Mao, A. Niewiadomska, M. Nguyen, G.J. Olsen, J.C. Overbeek, B. Parrello, V. Parrello, J.S. Porter, G.D. Pusch, M. Shukla, I. Singh, L. Stewart, G. Tan, C. Thomas, M. VanOeffelen, V. Vonstein, Z.S. Wallace, A.S. Warren, A.R. Wattam, F. Xia, H. Yoo, Y. Zhang,

C.M. Zmasek, R.H. Scheuermann, R.L. Stevens, Introducing the bacterial and viral bioinformatics resource center (BV-BRC): a resource combining PATRIC, IRD and ViPR, Nucleic Acids Res. 51 (2023) D678–D689, doi:10.1093/nar/gkac1003.

[14] J.R. Grant, E. Enns, E. Marinier, A. Mandal, E.K. Herman, C. Chen, M. Graham, G. Van Domselaar, P. Stothard, Proksee: in-depth characterization and visualization of bacterial genomes, Nucleic Acids Res. 51 (2023) W484–W492, doi:10.1093/nar/gkad326.

[15] T. Seemann, Abricate, Github (n.d.). https://github.com/tseemann/abricate.

[16] B.P. Alcock, A.R. Raphenya, T.T.Y. Lau, K.K. Tsang, M. Bouchard, A. Edalatmand, W. Huynh, A.-L.V Nguyen, A.A. Cheng, S. Liu, S.Y. Min, A. Miroshnichenko, H.-K. Tran, R.E. Werfalli, J.A. Nasir, M. Oloni, D.J. Speicher, A. Florescu, B. Singh, M. Faltyn, A. Hernandez-Koutoucheva, A.N. Sharma, E. Bordeleau, A.C. Pawlowski, H.L. Zubyk, D. Dooley, E. Griffiths, F. Maguire, G.L. Winsor, R.G. Beiko, F.S.L. Brinkman, W.W.L. Hsiao, G.V. Domselaar, A.G. McArthur, CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database, Nucleic Acids Res. 48 (D1) (2020) D517–D525, doi:10.1093/nar/gkz935.

[17] B. Liu, D. Zheng, S. Zhou, L. Chen, J. Yang, VFDB 2022: a general classification scheme for bacterial virulence factors, Nucleic Acids Res. 50 (2022) D912–D917, doi:10.1093/nar/gkab1107.

[18] J. Tabassum, A. Anjum, S. Islam, A. Khaleque, I. Jabeen, S.R. Shuvo, Draft genome sequencing data of multidrug-resistant *Staphylococcus haemolyticus* from Bangladeshi hospitals, Data Br. 57 (2024) 110918, doi:10.1016/j.dib.2024.110918.

[19] C. Jain, L.M. Rodriguez-R, A.M. Phillippy, K.T. Konstantinidis, S. Aluru, High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries, Nat. Commun. 9 (2018) 5114, doi:10.1038/s41467-018-07641-9.

[20] J. Sun, F. Lu, Y. Luo, L. Bie, L. Xu, Y. Wang, OrthoVenn3: an integrated platform for exploring and visualizing orthologous data across genomes, Nucleic Acids Res. 51 (2023) W397–W403, doi:10.1093/nar/gkad313.

[21] J.P. Meier-Kolthoff, M. Göker, TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy, Nat. Commun. 10 (2019) 2182, doi:10.1038/s41467-019-10210-3.

[22] R.S. Kaas, P. Leekitcharoenphon, F.M. Aarestrup, O. Lund, Solving the problem of comparing whole bacterial genomes across different sequencing platforms, PLoS One 9 (2014) e104984, doi:10.1371/journal.pone.0104984.

[23] N.M. Chaudhari, V.K. Gupta, C. Dutta, BPGA- an ultra-fast pan-genome analysis pipeline, Sci. Rep. 6 (2016) 24373, doi:10.1038/srep24373.