Title

Draft genome sequence of multidrug-resistant Escherichia coli MAHK_SCM_BAU_30A strain isolated from a subclinical mastitis cow in Bangladesh.

Journal

Microbiology Resource Announcements, 12(11)

Authors

Anika, Tasnia

Noman, Zakaria

Sultana, Nazneen

et al.

Publication Date

2023-11-16

DOI

10.1128/MRA.00713-23

Peer reviewed

AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

| Bacteriology | Announcement

Draft genome sequence of multidrug-resistant *Escherichia coli* MAHK_SCM_BAU_30A strain isolated from a subclinical mastitis cow in Bangladesh

Tasnia Tabassum Anika,[1] Zakaria Al Noman,[2] Md. Saiful Islam,[3,4] Nazneen Sultana,[1] Md. Nahid Ashraf,[3] Munmun Pervin,[1] Mohammad Ariful Islam,[5] Mokbul Md. Hossain,[1] Md. Tanvir Rahman,[3] Mohammad Abu Hadi Noor Ali Khan[1]

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT This study announces the sequence of a multidrug-resistant *Escherichia coli* MAHK_SCM_BAU_30A strain isolated from bovine subclinical mastitis milk in 2022 in Bangladesh. Our assembled genome had a length of 4,884,948 bp, three plasmids, two CRISPR arrays, five prophages, 51 predicted antibiotic resistance, and 72 predicted virulence factor genes.

KEYWORDS subclinical bovine mastitis, *E. coli*, whole genome sequencing, MDR, antibiotic resistance genes, virulence factor genes, CRISPR arrays, public health, Bangladesh

The global public health is at risk due to the widespread application and improper use of antibiotics, resulting in the emergence of antimicrobial resistance in various multidrug-resistant bacterial strains (1). Subclinical mastitis is a prevalent ailment in lactating dairy cows that may lead to diminished production and financial setbacks for farmers. *Escherichia coli* is recognized as the predominant bacterium capable of inducing subclinical mastitis and displaying resistance to antibiotics (2).

Between June and December 2022, milk samples were collected from cattle with subclinical mastitis in Baghabari (24.1369°N, 89.5859°E) within the Sirajganj district of Bangladesh and transported to the laboratory (24.7245°N, 90.4372°E). These samples were placed in nutrient broth (HiMedia, India), incubated at 37°C overnight, and subsequently spread on eosin methylene blue agar (HiMedia, India) media and incubated at 37°C overnight again. The resulting colonies underwent Gram staining and biochemical tests (indole, methyl red, Voges-Proskauer, citrate utilization, and sugar fermentation tests) to isolate *E. coli* (3). *E. coli* identification was accomplished using the matrix-assisted laser desorption ionization time-of-flight mass spectrometry assay (4). Antibiotic resistance was determined through the disk diffusion method (5) and the CLSI guidelines (6). Finally, a multidrug-resistant *E. coli* isolate, showing resistance to at least three antibiotic classes (7), was selected for this study and incubated overnight in nutrient broth (HiMedia, India) at 37°C. DNA was then extracted from the collected broth culture using a DNA mini kit (Qiagen, Hilden, Germany). The DNA concentration and purity were determined using a NanoDrop 2000 UV-Vis Spectrophotometer (Thermo Fisher, Waltham, MA, USA). The Nextera DNA Flex Library Prep Kit (Illumina, San Diego, CA, USA) was used to create the DNA library, and the genome sequencing was conducted on the Illumina NextSeq2000 platform, generating paired-end reads with a length of 2 × 150 bp. The genome assembly was performed using Unicycler v.0.4.9 (8), preceded by trimming the raw paired-end reads (n = 2,551,778) using Trimmomatic v.0.39 (9) (with parameters leading: 20, sliding window: 4:20:20, trailing: 20, minlen = 36), with the aim of eliminating Illumina adapters, recognized Illumina irregularities, and phiX reads from

Editor David Rasko, University of Maryland School of Medicine, Baltimore, Maryland, USA

Address correspondence to Mohammad Abu Hadi Noor Ali Khan, hadi.khan@bau.edu.bd.

The authors declare no conflict of interest.

See the funding table on p. 2.

Received 2 August 2023
Accepted 21 September 2023

Published 27 October 2023

the data set, and the quality was assessed using FastQC v.0.11.7 (10). The genome was annotated using PGAP v.3.0 (11). In our assembled genome, plasmids were predicted by PlasmidFinder v.2.1 (12); CRISPR arrays by CRISPRimmunity (13); prophages by PHASTER (14); antibiotic resistance genes (ARGs) by CARD v.3.2.4 (15) and ResFinder v.4.1 (16); virulence factor genes (VFGs) by VFDB (17) and VirulenceFinder v.2.0 (18); pathogenicity index by PathogenFinder v.1.1 (19); sequence type by MLST v.2.0 (20), and metabolic functional features by RAST v.2.0 (21). Default parameters were used for all tools unless specified otherwise.

The genome assembly of *E. coli* MAHK_SCM_BAU_30A strain comprised 128 contigs, featuring a G + C content of 50.69%. It included nine contig L50 with 161,852 bp of N_{50} value. The total genome size was 4,884,948 bp with a coverage of 16.35×. Within this genome, a total of 4,898 genes, 4,807 CDS, 78 tRNA genes, three rRNA genes, and 158 pseudogenes were identified. The genome contained two CRISPR arrays (with 10 genes, i.e., *csa3*, *cas2*, *cas1*, *cas6e*, *cas5*, *cas7*, *cse2gr11*, *cas8e*, *cas3*, and *WYL*), five prophages, and three plasmids [IncFIA, IncFIB, and IncFII(pHN7A8)]. Through MLST analysis, our genome was categorized as sequence type ST21, and the PlasmidFinder tool revealed a pathogenicity index of 0.945. Moreover, our genome harbored 51 predicted ARGs and 72 predicted VFGs. In RAST, 384 subsystems with 31% coverage and 2,158 genes were identified in our assembled genome.

ACKNOWLEDGMENTS

We would like to thank the Bangladesh Bureau of Educational Information and Statistics (BANBEIS), the Ministry of Education, the People's Republic of Bangladesh, for their financial support (Grant ID: LS20201343) which facilitated this study.

AUTHOR AFFILIATIONS

[1]Department of Pathology, Faculty of Veterinary Science, Bangladesh Agricultural University, Mymensingh, Bangladesh
[2]Bangladesh Council of Scientific and Industrial Research, Dhaka, Bangladesh
[3]Department of Microbiology and Hygiene, Faculty of Veterinary Science, Bangladesh Agricultural University, Mymensingh, Bangladesh
[4]Department of Animal Sciences, University of California—Davis, Davis, California, USA
[5]Department of Medicine, Faculty of Veterinary Science, Bangladesh Agricultural University, Mymensingh, Bangladesh

AUTHOR ORCIDs

Md. Saiful Islam http://orcid.org/0000-0002-6870-4595
Md. Tanvir Rahman http://orcid.org/0000-0001-5432-480X
Mohammad Abu Hadi Noor Ali Khan http://orcid.org/0000-0002-8409-4571

FUNDING

Funder	Grant(s)	Author(s)
Ministry of Education, Government of the People's Republic of Bangladesh (MoE)	LS20201343	Mohammad Abu Hadi Noor Ali Khan

AUTHOR CONTRIBUTIONS

Tasnia Tabassum Anika, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Writing – original draft | Zakaria Al Noman, Data curation, Investigation, Methodology | Md. Saiful Islam, Data curation, Formal analysis, Methodology, Software, Writing – original draft, Writing – review and editing | Nazneen Sultana, Investigation | Md. Nahid Ashraf, Investigation | Munmun Pervin, Methodology, Writing – review and editing | Mohammad Ariful Islam, Supervision | Mokbul Md. Hossain,

Supervision | Md. Tanvir Rahman, Conceptualization, Methodology, Validation, Writing – review and editing | Mohammad Abu Hadi Noor Ali Khan, Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Validation, Writing – review and editing

DATA AVAILABILITY

The WGS shotgun analysis of *E. coli* MAHK_SCM_BAU_30A was deposited to GenBank under the accession number JAUBOE000000000. The relevant data, including the raw reads, were also submitted with BioProject accession number PRJNA984821, BioSample accession number SAMN35779352, and SRA accession number SRR24954487. In this paper, the specific version being referred to is identified as JAUBOE000000000.1.

REFERENCES

1. Ahmed T, Islam MS, Nuruzzaman M, Sadekuzzaman M, Kabir SML, Rahman MT, Khan MSR. 2023. Draft genome sequence of the multidrug-resistant citrobacter freundii 132-2 strain isolated from a domestic duck in Bangladesh. Microbiol Resour Announc 12:e0037823. https://doi.org/10.1128/mra.00378-23
2. Hinthong W, Pumipuntu N, Santajit S, Kulpeanprasit S, Buranasinsup S, Sookrung N, Chaicumpa W, Aiumurai P, Indrawattana N. 2017. Detection and drug resistance profile of *Escherichia coli* from subclinical mastitis cows and water supply in dairy farms in Saraburi province, Thailand. PeerJ 5:e3431. https://doi.org/10.7717/peerj.3431
3. Hitchins AD, Feng P, Watkins WD, Rippey SR, Chandler LA, U.S. Food and Drug Administration. 1998. *Escherichia Coli* and the coliform bacteria, p 4. In Bacteriological analytical manual, 8th ed. AOAC International, Gaithersburg.
4. van Veen SQ, Claas ECJ, Kuijper EJ. 2010. High-throughput identification of bacteria and yeast by matrix-assisted laser desorption ionization-time of flight mass spectrometry in conventional medical microbiology laboratories. J Clin Microbiol 48:900–907. https://doi.org/10.1128/JCM.02071-09
5. Bauer AW, Kirby WM, Sherris JC, Turck M. 1966. Antibiotic susceptibility testing by a standardized single disk method. Am J Clin Pathol 45:493–496.
6. M100-S32. 2022. Performance standards for antimicrobial susceptibility testing. Clinical and Laboratory Standards Institute, Wayne, PA, USA.
7. Magiorakos AP, Srinivasan A, Carey RB, Carmeli Y, Falagas ME, Giske CG, Harbarth S, Hindler JF, Kahlmeter G, Olsson-Liljequist B, Paterson DL, Rice LB, Stelling J, Struelens MJ, Vatopoulos A, Weber JT, Monnet DL. 2012. Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. Clin Microbiol Infect 18:268–281. https://doi.org/10.1111/j.1469-0691.2011.03570.x
8. 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
9. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for illumina sequence data. Bioinf 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170
10. Andrews S. 2010. Online. *FastQC: a quality control tool for high throughput sequence data*. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
11. 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
12. 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
13. 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
14. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 44:W16–W21. https://doi.org/10.1093/nar/gkw387
15. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, Huynh W, Nguyen A-L, Cheng AA, Liu S, Min SY, Miroshnichenko A, Tran H-K, Werfalli RE, Nasir JA, Oloni M, Speicher DJ, Florescu A, Singh B, Faltyn M, Hernandez-Koutoucheva A, Sharma AN, Bordeleau E, Pawlowski AC, Zubyk HL, Dooley D, Griffiths E, Maguire F, Winsor GL, Beiko RG, Brinkman FSL, Hsiao WWL, Domselaar GV, McArthur AG. 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
16. Florensa AF, Kaas RS, Clausen P, Aytan-Aktug D, Aarestrup FM. 2022. Resfinder - an open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes. Microb Genom 8:000748. https://doi.org/10.1099/mgen.0.000748
17. 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
18. Kleinheinz KA, Joensen KG, Larsen MV. 2014. Applying the resfinder and virulencefinder web-services for easy identification of acquired antibiotic resistance and *E. Coli* virulence genes in bacteriophage and prophage nucleotide sequences. Bacteriophage 4:e27943. https://doi.org/10.4161/bact.27943
19. 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
20. Larsen MV, Cosentino S, Rasmussen S, Friis C, Hasman H, Marvig RL, Jelsbak L, Sicheritz-Pontén T, Ussery DW, Aarestrup FM, Lund O. 2012. Multilocus sequence typing of total-genome-sequenced bacteria. J Clin Microbiol 50:1355–1361. https://doi.org/10.1128/JCM.06094-11
21. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: rapid annotations using subsystems technology. BMC Genomics 9:75. https://doi.org/10.1186/1471-2164-9-75