BMC Research Notes

DATA NOTE Open Access

Genomic dataset of multidrug-resistant *Klebsiella quasipneumoniae* from Indonesia

Nastiti Intan Permata Sari[1], Putu Yuliandari[1], Ajeng Kusumaningtyas Pramono[1], Dwi Tyastuti[3], Abdul Hadi Furqoni[1], Fitriana Fitriana[1], Sunarno Sunarno[1], Dewi Setyowati[4*] and Ni Luh Putu Indi Dharmayanti[2]

Abstract

Objectives *Klebsiella pneumoniae* is a free-living bacterium found anywhere, including soil, water, and various types of plants, animals, and humans. Due to identical biochemical test results, *Klebsiella quasipneumoniae*, a member of the *Klebsiella pneumoniae* complex, is often misidentified as *Klebsiella pneumoniae*. This distinct species can be accurately identified solely through whole-genome sequencing. This bacterium poses a significant public health threat due to its increasing antibiotic resistance, ability to cause severe disease, and potential for community-acquired and hospital-acquired infections. However, there was no previous report of *K. quasipneumoniae* from Indonesia. Subsequent research focusing on antimicrobial-resistant gene analysis, virulence determinants, evolutionary relationship, and transmission pathways based on this dataset will enhance understanding of this species and their differences with other *Klebsiella pneumoniae* complex organisms.

Data description We present a whole genome sequencing of four *Klebsiella quasipneumoniae* isolated from hospital wastewater in Jakarta, Indonesia. Initial bacterial identification was conducted which showed *Klebsiella pneumoniae*. However, the whole genome Average Nucleotide Identity (wgANI) was found to be *Klebsiella quasipneumoniae*. The genome size of *Klebsiella quasipneumoniae* I1-1, I1-3, I5-2, and I5-3 isolates were 5.4 Mb (GC = 57.72%), 5.5 Mb (GC = 57.72%), 5.4 Mb (GC = 57.73%), and 5.5 Mb (GC = 57.72%), respectively. Sequence data has been deposited in the GenBank database.

Keywords *Klebsiella quasipneumoniae*, Indonesia, wastewater, whole-genome sequencing, public health

Objective

The genus Klebsiella belongs to the Enterobacteriaceae family, which is nonmotile, rod-shaped, and Gram-negative, and it produces a polysaccharide capsule. *K. pneumoniae* is a free-living bacterium found anywhere, including soil, water, and various types of plants, animals, and humans. This bacterium poses a significant public health threat due to its increasing antibiotic resistance, ability to cause severe disease, and potential for community-acquired and hospital-acquired infections [1, 2]. The *Klebsiella pneumoniae* complex is classified into three phylogroups: *Klebsiella pneumoniae* (KpI), *Klebsiella quasipneumoniae* (KpII), and *Klebsiella variicola* (KpIII)

*Correspondence:
Dewi Setyowati
dewi.setyowati@fk.unair.ac.id
[1]Center for Biomedical Research, Research Organization for Health, National Research and Innovation Agency (BRIN), Cibinong Science Centre, Bogor, West Java, Indonesia
[2]Research Organization for Health, National Research and Innovation Agency (BRIN), Cibinong Science Centre, Bogor, West Java, Indonesia
[3]Department of Community and Family Medicine, Faculty of Medicine, Universitas Islam Negeri Syarif Hidayatullah, Jakarta, Indonesia
[4]Midwifery Study Programme, Faculty of Medicine, Universitas Airlangga, Surabaya, Indonesia

Table 1 Overview of data files/data sets

Label	Name of data file/data sets	File types (file extension)	Data repository and identifier (DOI or accession number)
Data file 1	Data analysis workflow	PDF	Figshare: https://doi.org/10.6084/m9.figshare.28802672 [16]
Data file 2	Antimicrobial resistance profile	PDF	Figshare: https://doi.org/10.6084/m9.figshare.28805720 [17]
Data file 3	Klebsiella quasipneumoniae I1-1, I1-3, I5-2, and I5-3 NCBI BioProject	No file format	NCBI BioProject Database https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1206793 [18]
Data file 4	Klebsiella quasipneumoniae I1-1 Illumina raw sequence	fastq	NCBI Sequence Read Archive SRR31897982 https://identifiers.org/ncbi/insdc.sra:SRR31897982 [19]
Data file 5	Klebsiella quasipneumoniae I1-3 Illumina raw sequence	fastq	NCBI Sequence Read Archive SRR31897981 https://identifiers.org/ncbi/insdc.sra:SRR31897981 [20]
Data file 6	Klebsiella quasipneumoniae I5-2 Illumina raw sequence	fastq	NCBI Sequence Read Archive SRR31897980 https://identifiers.org/ncbi/insdc.sra:SRR31897980 [21]
Data file 7	Klebsiella quasipneumoniae I5-3 Illumina raw sequence	fastq	NCBI Sequence Read Archive SRR31897979 https://identifiers.org/ncbi/insdc.sra:SRR31897979 [22]
Data file 8	*Klebsiella quasipneumoniae* I1-1, I1-3, I5-2, and I5-3 NCBI BioSample	No file format	NCBI BioSample https://www.ncbi.nlm.nih.gov/biosample?LinkName=bioproject_biosample_all%26;from_uid=1206793 [23]

in classical taxonomy, but KpII and KpIII are rarely isolated clinically. *K. quasipneumoniae* is frequently mistaken as *K. pneumoniae* by conventional microbiology laboratories because of the resemblance in biochemical test findings. Whole-genome sequencing can help distinguish *K. quasipneumoniae* from *K. pneumoniae* and greatly increase the genetic understanding of *the K. quasipneumoniae* strain [3].

Wastewater surveillance has historically been used to track pathogens within a community. Compared with urban wastewater, hospital wastewater generally harbours substantial levels of antibiotic-resistant bacteria due to the intense selection pressure exerted by the frequent use of antibiotics. The presence of high multidrug-resistant (MDR) bacteria in hospital wastewater may impose public health challenges because they can transmit resistance traits to other enteric pathogenic bacteria in the community. Routine monitoring of hospital wastewater for pathogens and antimicrobial resistance genes may provide early detection and guide public health interventions [4].

This study investigates the genomic characteristics of four *Klebsiella quasipneumoniae* isolates obtained from hospital wastewater in Indonesia. To our knowledge, no previous report of *K. quasipneumoniae* from Indonesia existed. The data set of the pathogen bacteria presented here can provide valuable resources for genomic studies and future computational studies, including machine learning or deep learning models to understand genomic features or predict phenotypic traits related to *Klebsiella quasipneumoniae*.

Data description

Sampling sites and sample collection

This study was conducted on hospital wastewater in Jakarta, Indonesia, on 14th October 2024. The representative samples were taken using a "grab sampling technique" following the Indonesian Nasional Standard (SNI) 9063:2022 [5]. At the hospital wastewater treatment site, the covers of the manholes were carefully lifted, and 14 ml wastewater samples, each from the inlet and outlet site, were taken into sterile plastic containers. Samples were collected once at 1:00 P.M. All samples were collected manually and transported with a cold chain within two hours into the microbiology laboratory. Samples were stored in a refrigerator at 4^0C until they were processed.

Species identification test and antimicrobial susceptibility testing

Table 1 (Data file 1) presents a snapshot of our data analysis workflow. From 14 mL of each sample, 1 mL was taken to be diluted from 10^{-1} to 10^{-4} in sterile water. Then, 100 ul of each sample dilution was streaked onto MacConkey Agar (Oxoid, Cat. No. CM0007) using a sterile inoculating loop and incubated at 37^0C in the atmosphere for 24 h [6, 7]. A total of ten pink colonies (Gram-negative, lactose-fermenters) were taken randomly from each sample, inlet and outlet. The Clinical and Laboratory Standards Institute (CLSI) 2024 was used as an antimicrobial susceptibility testing reference [8–11]. Four colonies (I1-1, I1-3, I5-2, I5-3) were identified as *Klebsiella pneumoniae* and antimicrobial resistance results showed multidrug-resistance organisms (Table 1, Data File 2).

Whole genome sequencing (WGS) and draft genome assembly

Based on phenotypic results, four colonies identified as *Klebsiella pneumoniae* were processed for WGS library preparation. DNA was extracted using the Quick-DNA Magbead Plus Kit (D4082). The libraries were prepared

using the xGen DNA Library Prep EZ UNI Kit (IDT, 10009822). Sequencing was performed using 300 cycles (2 × 150 bp paired-end) of the Illumina sequencing reagent on the NextSeq 2000 platform. The raw reads were imported into Geneious Prime® 2024.0.3 for streamlined analysis (https://www.geneious.com). The reads were filtered based on base quality (> Q30) and length (> 50 bp) using BBDuk, a tool included in the BBMap package (https://sourceforge.net/projects/bbmap/). Next, the filtered reads were paired and assembled using SPAdes Assembler 3.15.5 [12], and only contigs longer than 1,000 bp were retained for further analysis. Annotation of coding regions, RNA genes, and other genomic features was done through the RAST-tk pipeline [13].

Draft genomes were constructed from contigs using the Mauve MCM algorithm [14] and compared with representative genomes of the Klebsiella spp. in the database. This process used whole-genome Average Nucleotide Identity (wgANI) using FastANI26 v.1.34 by the default setting [15]. The species with the highest wgANI was taxonomically assigned to the assembled genome. For species identification, the 95% identity criteria were considered. There were *Klebsiella quasipneumoniae* I1-1 (98.7%), *Klebsiella quasipneumoniae* I1-3 (99.2%), *Klebsiella quasipneumoniae* I5-2 (98.7%); and *Klebsiella quasipneumoniae* I5-3 (98.7%) with *Klebsiella quasipneumoniae* subsp. quasipneumoniae strain KP_NORM_BLD_60803; COP15399 as reference. An overview of the genomic dataset is presented in Table 1 (Data files 3–8).

Limitations

This data note focuses on the genomic characteristics of multidrug-resistant *Klebsiella quasipneumoniae* isolates obtained from hospital wastewater. This study has several limitations, including a small sample size (only four isolates) and a limited scope (single hospital wastewater site). More in-depth research is needed to understand phylogenetics, antimicrobial-resistant gene analysis, virulence factors, and one health approach to investigate transmission routes.

Acknowledgements
None.

Author contributions
Wrote the manuscript; (N.I.P.S., P.Y., A.K.P., A.H.F., F.F., S.S., D.S.,): Designed the study; (N.I.P.S., P.Y., A.K.P.,): investigation; (N.I.P.S., P.Y., A.K.P., D.S., N.L.P.I.D.,): conducted in silico analyses and generated the results; (P.Y., A.K.P.,): critically reviewed and revised the manuscript; (P.Y., N.I.P.S.,): supervised the study; (N.I.P.S., D.S.,): Resources; (A.H.F., D.T.,).

Funding
This study received no external funding.

Data availability
The data from this study has been deposited under BioProject PRJNA1206793. The raw reads obtained from the Illumina NextSeq 2000 sequencer have been submitted to the NCBI SRA database (SRS23701602; SRS23701603; SRS23701604; SRS23701605) and are publicly accessible. The data access links for all the data mentioned above are provided in Table 1.

Declarations

Ethics approval and consent to participate
This research was granted ethical clearance and approved by the Ethical Committee in the Health Research Unit of UIN Syarif Hidayatullah Jakarta, Indonesia, which obtained written No. B-024/F12/KEPK/TL.00/EE-10/2024.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 1 February 2025 / Accepted: 25 April 2025
Published online: 15 May 2025

References

1. Iwase T, Ogura Y, Hayashi T, Mizunoe Y. Complete genome sequence of Klebsiella pneumoniae YH43. Genome Announc. 2016;4(2):e00242–16. https://doi.org/10.1128/genomea.00242-16.
2. Fatmawati NND, Tarini NMA, Budayanti NNS, Yuliandari P. Molecular characterization of Extended-spectrum b-lactamase-producing *Lebsiellapneumoniae* isolated from clinical specimens at a tertiary-referral hospital in Denpasar, Bali, Indonesia. Adv Sci Lett. 2015;21:219–21. https://doi.org/10.1166/asl.2015.5860.
3. Liu Y, Liu X, Xu H, Zhang X, Liu R, Chen M, Qian J, Zheng B. Genomic and phenotypic characterization of ST2012 clinical Klebsiella quasipneumoniae subsp. Similipneumoniae harboring blaNDM– 1 in China. BMC Microbiol. 2024;24:506. https://doi.org/10.1186/s12866-024-03637-2.
4. Akinola OT, Dahunsi SO. Whole genome sequencing reveals antibiotic resistance and virulence factors in Klebsiella quasipneumoniae subsp. Similipneumoniae from hospital wastewater in South-West. Nigeria Microb Pathog. 2024;197:107040. https://doi.org/10.1016/j.micpath.2024.107040.
5. Standar Nasional Indonesia (SNI). 9063:2022 Tentang metode Pengambilan Contoh Uji air Dan air Limbah Untuk parameter mikrobiologi. Jakarta: Badan Standardisasi Nasional; 2022.
6. Gajic I, Kabic J, Kekic D, Jovicevic M, Milenkovic M, Culafic DM, Trudic A, Ranin L, Opavski N. Antimicrobial susceptibility testing: A comprehensive review of currently used methods. Antibiot (Basel). 2022;11(4):1–26. https://doi.org/10.3390/antibiotics11040427.
7. Tuttle AR, Trahan ND, Son MS. Growth and maintenance of Escherichia coli laboratory strains. Curr Protoc. 2021;1(1):e20. https://doi.org/10.1002/cpz1.20.
8. CLSI. 2024. Performance Standards for Antimicrobial Susceptibility Testing. 34th ed. CLSI Supplement M100. Clinical and Laboratory Standards Institute, USA.
9. Morris S, Cerceo E. Trends, epidemiology, and management of multidrug-resistant gram-negative bacterial infections in the hospitalized setting. Antibiot (Basel). 2020;9(4):196. https://doi.org/10.3390/antibiotics9040196.
10. Baswendra T, Suwarno S, Sarudji S, Damayanti R, Sugihartuti R, Estoepangesti ATS. Antibiotic sensitivity test of Escherichia coli and Staphylococcus aureus isolated from the reproductive tract of dairy cows. Ovozoa: J Anim Reprod. 2022;11(2):72–80. https://doi.org/10.20473/ovz.v11i2.2022.72-80.
11. Khan S, Das A, Mishra A, Vidyarthi AJ, Nandal M, Yadav H, Roy S, Singh M. Evaluation of three protocols for direct susceptibility testing for Gram-negative rods from flagged positive blood culture bottles. Microbiol Spectr. 2024;12(4):e0308123. https://doi.org/10.1128/spectrum.03081-23.
12. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19(5):455–77. https://doi.org/10.1089/cmb.2012.0021.
13. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, 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. The RAST

server: rapid annotations using subsystems technology. BMC Genomics. 2008;9(1):75. https://doi.org/10.1186/1471-2164-9-75.
14. Rissman AI, Mau B, Biehl BS, Darling AE, Glasner JD, Perna NT. Reordering contigs of draft genomes using the mauve aligner. Bioinformatics. 2009;25(16):2071–3. https://doi.org/10.1093/bioinformatics/btp356.
15. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun. 2018;9:5114. https://doi.org/10.1038/s41467-018-07641-9.
16. Yuliandari P. Data file 1. Data analysis workflow used for Whole-genome Sequencing (WGS) of Bacterial Isolate. https://doi.org/10.6084/m9.figshare.28802672. In. Figshare; 2025.
17. Yuliandari P. Data file 2. Antimicrobial resistance profile. https://doi.org/10.6084/m9.figshare.28805720. In. Figshare; 2025.
18. Pramono AK. Data file 3. NCBI bioproject database. In. NCBI; 2025. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1206793.
19. Pramono AK. Data file 4. NCBI Sequence Read Archive (SRR31897982). https://identifiers.org/ncbi/insdc.sra:SRR31897982. In. NCBI-SRA; 2025.
20. Pramono AK. Data file 5. NCBI Sequence Read Archive (SRR31897981). https://identifiers.org/ncbi/insdc.sra:SRR31897981. In. NCBI-SRA; 2025.
21. Pramono AK. Data file 6. NCBI Sequence Read Archive (SRR31897980). https://identifiers.org/ncbi/insdc.sra:SRR31897980. In. NCBI-SRA; 2025.
22. Pramono AK. Data file 7. NCBI Sequence Read Archive (SRR31897979). https://identifiers.org/ncbi/insdc.sra:SRR31897979. In. NCBI-SRA; 2025.
23. Pramono AK. Data file 8. NCBI biosample database. biosample?LinkName=bioproject_biosample_all&from_uid=1206793 In. NCBI; 2025. https://www.ncbi.nlm.nih.gov/.

Publisher's note