

Complete genome sequence of the multidrug-resistant *Citrobacter portucalensis* strain KOS1-1 isolated from a wastewater treatment plant

Shahjahon Begmatov,¹ Andrey L. Rakitin,¹ Alexey V. Beletsky,¹ Andrey V. Mardanov,¹ Nikolai V. Ravin¹

AUTHOR AFFILIATION See affiliation list on p. 2.

ABSTRACT The emergence of multidrug-resistant bacteria is a major concern from both epidemiological and ecological perspectives. Studying these microorganisms is crucial for developing prevention strategies. We report the complete genome sequence of the multidrug-resistant *Citrobacter portucalensis* strain KOS1-1, isolated from wastewater in the city of Moscow.

KEYWORDS multidrug-resistant bacteria, *Citrobacter*, plasmid-mediated resistance, wastewater treatment

The growing interest in multidrug-resistant (MDR) strains from wastewater treatment plants (WWTP) stems from their role as potential reservoirs of mobile antibiotic resistance genes (ARGs). Consequently, their genetic makeup and resistomes are increasingly being characterized. MDR *Citrobacter* strains have been isolated from WWTP-related samples, and many of them produce carbapenemase and extended-spectrum β -lactamase, and these genes are consistently associated with a wide spectrum of mobile genetic elements (1). MDR bacteria can survive wastewater treatment processes and be released into the environment with the effluent (2). Moscow's WWTPs, among the largest in the world, process approximately 6×10^6 m³ of water per day. Treated wastewater is discharged into the Moscow River, constituting nearly 50% of its total flow. Due to their immense scale, these WWTPs represent an important source for the dissemination of ARGs into the environment (3).

C. portucalensis strain KOS1-1 was isolated by plating of untreated influent wastewater on LB agar medium supplemented with tetracycline (20 μ g/mL), chloramphenicol (20 μ g/mL), kanamycin (50 μ g/mL), and ampicillin (50 μ g/mL) (4). The strain was grown in LB broth supplemented with these antibiotics at 20°C for 2 days; genomic DNA was extracted from 50 mL of culture using the DNeasy PowerSoil Kit (Qiagen, Germany) according to the manufacturer's protocol.

A hybrid sequencing approach combining Illumina and Oxford Nanopore platforms was used to sequence the genome. The library for Illumina sequencing was prepared using the NEBNext Ultra II DNA Library prep kit (New England Biolabs, USA). Its sequencing on the MiSeq instrument (Illumina, USA) generated 741,626 read pairs (2 $\times$ 300 nt) with a total length of 446.45 Mb. Overlapping reads were merged using FLASH v.1.2.11 (5) with "M 300" parameter, low-quality sequences were trimmed with Sickle v.1.33 (<https://github.com/najoshi/sickle>) with a quality threshold of Q₃₀. Additionally, the *C. portucalensis* KOS1-1 genome library was prepared from non-sheared DNA using the 1D Genomic DNA ligation sequencing kit SQK-LSK114 and sequenced in a FLO-MIN114 flow cell using a MinION device (Oxford Nanopore, UK). A total of 17,354 reads with a total length of 276 Mb were obtained. The N₅₀ value was 23,343 bp, and the average

Editor André O. Hudson, Rochester Institute of Technology, Rochester, New York, USA

Address correspondence to Shahjahon Begmatov, shahbegmatov@gmail.com.

The authors declare no conflict of interest.

See the funding table on p. 3.

See the related article at <https://doi.org/10.1128/spectrum.02038-25>.

Received 3 December 2025

Accepted 28 January 2026

Published 19 February 2026

Copyright © 2026 Begmatov et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](#).

TABLE 1 Genome sequencing statistics

Parameter	Value
Illumina sequencing	
Read pairs number	741,626
Total bases	446,458,852
Nanopore sequencing	
Reads number	18,901
Total bases	298,186,162
Genome assembly	
Total length, bp	5,556,809
Genome coverage	99.0
Chromosome, bp	4,878,312
Plasmid pKOS1-1-1, bp	77,569
Plasmid pKOS1-1-2, bp	80,667
Plasmid pKOS1-1-3, bp	110,147
Plasmid pKOS1-1-4, bp	121,081
Plasmid pKOS1-1-5, bp	289,033
GC content (%)	51.69
Protein-coding genes	5,227
tRNA genes	85
rRNA genes	25
ncRNA genes	7
Pseudo genes	178

read length was 15,880 bp. Basecalling was computed by Dorado v.0.9.1 (6) using the dna_r10.4.1_e8.2_400bps_sup@v5.0.0 model.

The MinION reads were assembled using Flye v. 2.9.5 (7) with --nano-hq and --iterations 2 parameters. The sequences of the assembled contigs were polished with Illumina reads using Pypolca v.0.3.1 (8). The assembled complete genome was 5,556,809 bp in size and consisted of a chromosome and five circular plasmids. Genes were predicted using NCBI Prokaryotic Genome Annotation Pipeline (version 2023-05-17.build6771) (9). Approximately 99% of Illumina and 88% of MinION reads passed quality filters. Genome statistics are described in Table 1.

The average nucleotide sequence identity between the genomes of *C. portucalensis* KOS1-1 and the type strain *C. portucalensis* A60^T (10), calculated by FastANI v.1.33 (11), was 98.28%, confirming the species assignment of strain KOS1-1. All tools were run with default parameters unless otherwise specified.

The genome contained 14 ARGs that were predicted to confer resistance to β -lactams, quinolones, aminoglycosides, macrolides, sulfonamides, trimethoprim, phenicols, and tetracyclines (4). The complete genome of *C. portucalensis* KOS1-1 facilitates comparative genomic studies of MDR *Citrobacter* strains.

ACKNOWLEDGMENTS

This study was performed using the scientific equipment of the Core Research Facility "Bioengineering" (Research Center of Biotechnology RAS) and was partly funded by the Russian Science Foundation (project no. 24-74-10045 to S.B.).

AUTHOR AFFILIATION

¹Institute of Bioengineering, Research Center of Biotechnology of the Russian Academy of Sciences, Moscow, Russia

AUTHOR ORCIDs

Shahjahon Begmatov  <http://orcid.org/0000-0002-3967-9247>

Andrey V. Mardanov  <http://orcid.org/0000-0002-8245-8757>

Nikolai V. Ravin  <http://orcid.org/0000-0002-1456-1832>

FUNDING

Funder	Grant(s)	Author(s)
Russian Science Foundation	24-74-10045	Shahjahon Begmatov

AUTHOR CONTRIBUTIONS

Shahjahon Begmatov, Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Writing – original draft | Andrey L. Rakitin, Investigation, Methodology | Alexey V. Beletsky, Investigation, Methodology, Writing – original draft | Andrey V. Mardanov, Data curation, Investigation, Methodology | Nikolai V. Ravin, Conceptualization, Supervision, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

The whole-genome nucleotide sequence of *C. portucalensis* KOS1-1 has been deposited in DDBJ/ENA/GenBank under accession numbers [CP199069](#), [CP199070](#), [CP199071](#), [CP199072](#), [CP199073](#), and [CP199074](#), and the raw reads have been deposited in the NCBI Sequence Read Archive under accession numbers [SRX31239522](#) and [SRX31239521](#). The version described in this paper is the first version.

REFERENCES

- Hamerlinck H, Aerssens A, Boelens J, Dehaene A, McMahon M, Messiaen AS, Vandendriessche S, Velghe A, Leroux-Roels I, Verhasselt B. 2023. Sanitary installations and wastewater plumbing as reservoir for the long-term circulation and transmission of carbapenemase producing *Citrobacter freundii* clones in a hospital setting. *Antimicrob Resist Infect Control* 12:58. <https://doi.org/10.1186/s13756-023-01261-9>
- Sambaza SS, Naicker N. 2023. Contribution of wastewater to antimicrobial resistance: a review article. *J Glob Antimicrob Resist* 34:23–29. <https://doi.org/10.1016/j.jgar.2023.05.010>
- Begmatov S, Beletsky AV, Dorofeev AG, Pimenov NV, Mardanov AV, Ravin NV. 2024. Metagenomic insights into the wastewater resistome before and after purification at large-scale wastewater treatment plants in the Moscow city. *Sci Rep* 14:6349. <https://doi.org/10.1038/s41598-024-56870-0>
- Begmatov S, Rakitin AL, Beletsky AV, Mardanov AV, Ravin NV. 2026. Plasmids of the multidrug-resistant *Citrobacter portucalensis* KOS1-1 strain isolated from a wastewater treatment plant harbor antibiotic resistance genes and gene clusters involved in carbon metabolism. *Microbiol Spectr*:e0203825. <https://doi.org/10.1128/spectrum.02038-25>
- Magoč T, Salzberg SL. 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27:2957–2963. <https://doi.org/10.1093/bioinformatics/btr507>
- Oxford Nanopore Technologies PLC. 2025. GitHub - Nanoporetech/dorado: Oxford Nanopore's Basecaller. <https://github.com/nanoporetech/dorado/>.
- 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>
- Zimin AV, Salzberg SL. 2020. The genome polishing tool POLCA makes fast and accurate corrections in genome assemblies. *PLoS Comput Biol* 16:e1007981. <https://doi.org/10.1371/journal.pcbi.1007981>
- Li W, O'Neill KR, Haft DH, DiCuccio M, Chetvernin V, Badretin A, Coulouris G, Chitsaz F, Derbyshire MK, Durkin AS, Gonzales NR, Gwadz M, Lanczycki CJ, Song JS, Thanki N, Wang J, Yamashita RA, Yang M, Zheng C, Marchler-Bauer A, Thibaud-Nissen F. 2021. RefSeq: expanding the prokaryotic genome annotation pipeline reach with protein family model curation. *Nucleic Acids Res* 49:D1020–D1028. <https://doi.org/10.1093/nar/gkaa1105>
- Ribeiro TG, Gonçalves BR, da Silva MS, Novais Â, Machado E, Carriço JA, Peixe L. 2017. *Citrobacter portucalensis* sp. nov., isolated from an aquatic sample. *Int J Syst Evol Microbiol* 67:3513–3517. <https://doi.org/10.1099/ijsem.0.002154>
- Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 9:5114. <https://doi.org/10.1038/s41467-018-07641-9>