

Chromosome-scale scaffold genome sequence of *Arthrobacter citreus* strain NJF-O2 isolated from the distal colon of a healthy sheep

A. S. Wadday,^{1,2} Baraa Akeel Al-Hasan,^{2,3,4,5} Ali H. D. Janabi³

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT *Arthrobacter citreus* strain NJF-O2 was isolated from the distal colon of a healthy adult ram postmortem in Najaf Governorate, Iraq. We present the chromosome-scale scaffold genome sequence of this strain to expand *Arthrobacter* genomic resources and enable high-resolution comparative and functional genomic studies of its metabolic plasticity and environmental adaptability.

KEYWORDS *Arthrobacter citreus*, chromosome-scale scaffold genome, ruminant microbiota

Arthrobacter species are gram-positive, high-guanine-cytosine (G + C) Actinobacteria commonly found in soil, rhizospheres, and other environments renowned for their metabolic diversity, stress tolerance, and bioremediation potential (1). The genus *Arthrobacter* has undergone substantial taxonomic revision based on phylogenetic and genomic evidence, and chromosome-scale scaffold genome sequences provide stable references for comparative analyses within the genus (2). This chromosome-scale scaffold expands genomic resources for *Arthrobacter* and provides a useful reference for comparative analyses of genome architecture, gene context, and ecological diversity within the genus. *Arthrobacter citreus* strain NJF-O2 was isolated from the distal colon of a clinically healthy adult ram obtained postmortem at a licensed slaughterhouse in Najaf Governorate, Iraq. Approximately 1 g of luminal content and multiple mucosal swabs were collected using sterile swabs, placed in sterile containers on ice, and transported promptly to the laboratory. Upon receipt, aliquots were prepared as 10-fold serial dilutions in sterile phosphate-buffered saline and plated on Brucella blood agar, clarified rumen fluid agar, and brain heart infusion agar (Thermo Fisher Scientific, USA). Plates were incubated at 37°C for 48 h in a candle jar under reduced-oxygen conditions (~3–5% O₂). For whole-genome sequencing, a loopful of purified culture was applied to an ANICARD (SAN Group Biotech Germany GmbH) and shipped under sterile conditions to SAN Group Laboratories (Germany) for DNA extraction and quantification. Genomic DNA was extracted from a pure culture using a commercial kit. DNA quality and purity were assessed spectrophotometrically (UV-1900i Plus, Shimadzu), and samples with A260/280 ratios of 1.8–2.0 were used for sequencing, and sequencing libraries were prepared using the Watchmaker DNA Library Prep Kit (Watchmaker Genomics) and sequenced on the Illumina MiSeq platform (2 × 151 bp paired-end reads), yielding approximately 3.3 million read pairs and an estimated genome coverage of 184x. Reads were quality-trimmed using Trimmomatic v0.39 (3) and *de novo* assembled with SPAdes v3.15.3 (4). Assembly statistics calculated with QUAST v5.2.0 (5) showed that the initial assembly comprised 14 contigs, with an N50 of 875,329 bp and a largest contig of 1,101,369 bp. GTDB-Tk v2.4.0, data release r220 (6), classified the genome as *Arthrobacter_B citreus* based on average nucleotide identity (ANI) screening against reference genome

Editor Julie C. Dunning Hotopp, University of Maryland School of Medicine, Baltimore, Maryland, USA

Address correspondence to A. S. Wadday, ali.wadday@iunajaf.edu.iq.

The authors declare no conflict of interest.

Received 18 March 2026

Accepted 2 April 2026

Published 30 April 2026

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

[GCF_009192745.1](#) (ANI, 98.97%; alignment fraction, 0.958). The draft contigs were first scaffolded with MeDuSa v1.6 (7) using multiple closely related *Arthrobacter* reference genomes (Table S1; available at <https://doi.org/10.6084/m9.figshare.31796188>), yielding two scaffolds. The resulting assembly was then refined with RagTag v2.1.0 (8) using the closely related reference genome GCF_029023805.1 (*A. citreus* strain DSM 15046), producing a chromosome-scale scaffold. Genome completeness and contamination were assessed using CheckM v1.2.3 (9), which estimated the assembly to be 98.39% completeness with 0.98% contamination. The scaffold was reoriented to begin at the dnaA gene using Circlator's fixstart module (10). Assembly consistency was assessed by comparing terminal scaffold regions using SeqKit v2.6.1 (11) and NUCmer (MUMmer v4.0.0) (12). Default parameters were used for all software, unless otherwise noted. Because the final assembly structure was informed by reference-guided scaffolding, the genome is conservatively described here as a chromosome-scale scaffold. The chromosome-scale scaffold comprises 3,527,331 bp with a G + C content of 65.63%. Annotation via the National Center for Biotechnology Information (NCBI) Prokaryotic Genome Annotation Pipeline (PGAP v6.10) (13) predicted 3,300 genes, 3,239 protein-coding sequences, 52 tRNA genes, one complete rRNA operon (5S, 16S, 23S), four noncoding RNAs, and 28 pseudogenes. This chromosome-scale scaffold genome sequence of *A. citreus* strain NJF-O2 provides a useful genomic resource for future comparative and functional studies.

ACKNOWLEDGMENTS

We thank SAN Group Laboratories (Germany) for technical assistance with whole-genome sequencing.

AUTHOR AFFILIATIONS

¹Department of Anesthesia Techniques, College of Medical Technology, The Islamic University, Najaf, Iraq

²Iraqi Biosafety Association, Najaf, Iraq

³Department of Veterinary Microbiology, College of Veterinary Medicine, University of Al-Qadisiyah, Al-Diwaniyah, Iraq

⁴Department of Medical Laboratory Technology, College of Medical Technology, The Islamic University, Najaf, Iraq

⁵Department of Laboratory, Al Najaf Veterinary Hospital, Najaf, Iraq

AUTHOR ORCID*s*

A. S. Wadday  <http://orcid.org/0009-0001-7532-5243>

Baraa Akeel Al-Hasan  <http://orcid.org/0000-0001-6932-727X>

Ali H. D. Janabi  <http://orcid.org/0000-0002-0890-4894>

AUTHOR CONTRIBUTIONS

A. S. Wadday, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review and editing | Baraa Akeel Al-Hasan, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review and editing | Ali H. D. Janabi, Project administration, Supervision

DATA AVAILABILITY

The chromosome-scale scaffold genome sequence of *Arthrobacter citreus* strain NJF-O2 has been deposited in GenBank under accession [CP197463](#). The associated BioProject accession number is [PRJNA1251031](#), and the BioSample accession number is [SAMN47990684](#). Raw sequencing reads have been deposited in the Sequence Read Archive under accession [SRR36990657](#).

REFERENCES

1. Mongodin EF, Shapir N, Daugherty SC, DeBoy RT, Emerson JB, Shvartzbeyn A, Radune D, Vamathevan J, Riggs F, Grinberg V, Khouri H, Wackett LP, Nelson KE, Sadowsky MJ. 2006. Secrets of soil survival revealed by the genome sequence of *Arthrobacter aurescens* TC1. *PLoS Genet* 2:e214. <https://doi.org/10.1371/journal.pgen.0020214>
2. Busse H-J. 2016. Review of the taxonomy of the genus *Arthrobacter*, emendation of the genus *Arthrobacter* sensu lato, proposal to reclassify selected species of the genus *Arthrobacter* in the novel genera *Glutamicibacter* gen. nov., *Paeniglutamicibacter* gen. nov., *Pseudoglutamicibacter* gen. nov., *Paenarthrobacter* gen. nov. and *Pseudarthrobacter* gen. nov., and emended description of *Arthrobacter roseus*. *Int J Syst Evol Microbiol* 66:9–37. <https://doi.org/10.1099/ijsem.0.000702>
3. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
4. 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. 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>
5. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>
6. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2020. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* 36:1925–1927. <https://doi.org/10.1093/bioinformatics/bt z848>
7. Bosi E, Donati B, Galardini M, Brunetti S, Sagot M-F, Lió P, Crescenzi P, Fani R, Fondi M. 2015. MeDuSa: a multi-draft based scaffold. *Bioinformatics* 31:2443–2451. <https://doi.org/10.1093/bioinformatics/bt v171>
8. Alonge M, Leonard AS. 2021. Malonge/RagTag: RagTag v2.1.0. Zenodo.
9. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>
10. Hunt M, Silva ND, Otto TD, Parkhill J, Keane JA, Harris SR. 2015. Circlator: automated circularization of genome assemblies using long sequencing reads. *Genome Biol* 16:294. <https://doi.org/10.1186/s13059-015-0849-0>
11. Shen W, Le S, Li Y, Hu F. 2016. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One* 11:e0163962. <https://doi.org/10.1371/journal.pone.0163962>
12. Marçais G, Delcher AL, Phillippy AM, Coston R, Salzberg SL, Zimin A. 2018. MUMmer4: a fast and versatile genome alignment system. *PLoS Comput Biol* 14:e1005944. <https://doi.org/10.1371/journal.pcbi.1005944>
13. 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>