



Complete genome sequence of *Rhodococcus erythropolis* strain ICBD2, isolated from hydrocarbon-polluted coastal sediments in Central Chile

Ximena Báez-Matus,^{1,2} Roberto E. Durán,^{1,2} Roberto Orellana,^{2,3,4} Andrés Cumsille,¹ Ester G. Rivera,¹ Teresa Esparza-Correa,^{1,2} Francisco Salvà-Serra,^{5,6,7,8} Daniel Jaén-Luchoro,^{5,6,7} Edward R. B. Moore,^{5,6,7} Michael Seeger^{1,2}

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT *Rhodococcus erythropolis* strain ICBD2 (=CCUG 74281) is a hydrocarbon-degrading bacterium isolated from chronically hydrocarbon-polluted saline coastal sediments from the Valparaíso Region, Central Chile. The complete genome sequence of *R. erythropolis* ICBD2 (62.3% GC content) comprises a 6.35 Mb circular chromosome and a 387 kb linear plasmid, with 6,196 coding sequences.

KEYWORDS *Rhodococcus erythropolis*, genome, hydrocarbon, Central Chile

Rhodococcus genus (*Actinomycetota* phylum) comprises aerobic, non-motile, metabolically versatile Gram-positive bacteria that degrade persistent organic pollutants and synthesize antimicrobial molecules (1–4). Environmental bacteria possess diverse adaptation mechanisms to oxidative and saline stress (5–7). This study presents the complete genome sequence of the hydrocarbon-degrading bacterium *Rhodococcus erythropolis* ICBD2 (=CCUG 74281), isolated from a chronically hydrocarbon-polluted saline coastal sediment at the Las Salinas site (33.0011° S, 71.5469° W) in Viña del Mar (Valparaíso Region, Central Chile). Strain ICBD2 was isolated in 2016 after two successive enrichments of a sediment sample (1 g, 4 m depth) in Bushnell–Haas (BH) mineral medium (20 mL) using diesel (1% vol/vol) and acetate (10 mM) as sole carbon/energy sources without agitation. Thereafter, the enrichment culture without prior dilution (100 µL) was plated onto BH agar containing acetate (10 mM) as sole carbon/energy sources; a single colony was isolated and characterized. Strain ICBD2 grew in BH or low-phosphate Tris-buffered mineral salt media on *n*-dodecane, *n*-tetradecane, *n*-hexadecane, *n*-eicosane, and diesel as sole carbon/energy sources (30°C).

Genomic DNA was extracted from a single colony grown on tryptic soy agar using a modified Marmur procedure (8). An Illumina standard genomic paired-end library, prepared from DNA using an optimized protocol (Eurofins Genomics, Konstanz, Germany) with standard Illumina adapters, was Illumina-sequenced (NovaSeq-6000-platform/flow-cell S2), achieving 22,030,486 paired-end reads/151 bp average length (3,326 Mb). Quality was assessed using FastQC (v0.11.9) (9). DNA was complementary sequenced using the Oxford Nanopore Technologies MinION Mk101B sequencer. The library was prepared with DNA without pre-treatment and Rapid-Barcoding-Kit (vR9) (SQK-RBK004) and Nanopore-sequenced (48 h) using a flow cell FLO-MIN106 (vR9.4). Base calling was done (Guppy, v4.0.14), obtaining 157,563 long reads/10,937 bp average length. The hybrid assembly using SPAdes (v3.13.1) with read error correction (10) yielded a single circular chromosome (6,351,314 bp/62.3% GC) and a linear plasmid (387,366 bp/61.1% GC), containing inverted repeats at the ends. The overlaps were identified by BLASTN analysis of the ends against the genome sequence. The overlap of one of the chromosome ends was trimmed manually using UGENE (v36). The *dnaA* gene

Editor Silvia T. Cardona, University of Manitoba, Winnipeg, Manitoba, Canada

Address correspondence to Michael Seeger, michael.seeger@usm.cl.

The authors declare no conflict of interest.

See the funding table on p. 3.

Received 31 March 2026

Accepted 11 May 2026

Published 29 May 2026

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

located by TBLASTN analysis of the Unicycler (v0.4.8) start gene database was chosen as the start gene to re-orient the chromosome (UGENE v36). Genome completeness (CheckM2, v1.0.2) was assessed (11). Genes were annotated using the NCBI Prokaryotic Genome Annotation Pipeline (v6.10) (12). All software was utilized with default parameters unless otherwise noted. The ICBD2 genome (100% completeness) contains 6,196 coding sequences, 54 tRNA genes, and 15 rRNA genes.

Comparative whole-genome sequence analyses with the Type Strain Genome Server platform (13) and Pyani-plus (v1.0.0) (14) identified strain ICBD2 as *Rhodococcus erythropolis*; 16S rRNA gene identity (100%), digital DNA–DNA hybridization ($d_4 = 89.3\%$), and average nucleotide identity (ANIb = 98.5%) with *Rhodococcus erythropolis* NBRC 15567^T.

BLAST gene function analyses using the UniProt-KB/Swiss-Prot database (>30% identity) identified genes associated with alkane degradation and oxidative stress response, including five *alkB* genes encoding alkane monooxygenase ($\geq 69\%$ identity with *Rhodococcus* spp. *alkB* genes, ACV96D_03880/ACV96D_10690/ACV96D_26645/ACV96D_11635/ACV96D_31060), *katG* gene encoding catalase-peroxidase (ACV96D_05890), *katA* (ACV96D_02200), and *katE* (ACV96D_28135) genes encoding catalases, and *sod* genes encoding superoxide dismutase ([Mn]SodA/ACV96D_00775, [Fe]SodB/ACV96D_23855, [Cu-Zn]SodC/ACV96D_07375). The *Rhodococcus erythropolis* ICBD2 genetic repertoire highlights its potential for hydrocarbon biodegradation.

ACKNOWLEDGMENTS

The authors acknowledge Timur Tunovic for DNA extraction and Las Salinas S.A. for providing access to the Las Salinas site and facilitating sediment sampling in the frame of a collaborative agreement between Las Salinas S.A. and Universidad Técnica Federico Santa María, aimed at promoting scientific research on environmental remediation and the sustainable management of impacted industrial sites.

This study was supported by ANID PhD scholarships (numbers 21211723 [X.B.-M.] and 21242707 [R.E.D.], PUCV [X.B.-M.], USM [X.B.-M.], and PIIC 055/2021 [X.B.-M.]); PhD scholarships; ANID Master scholarship (2230193 [T.E.-C.]); and Agencia Nacional de Investigación y Desarrollo ANID-Milenio-NCN2023_054 (M.S., X.B.-M., R.E.D., T.E.-C., and R.O.), Fondo Nacional de Desarrollo Científico y Tecnológico Fondecyt (1250809 and 1200756 [M.S., X.B.-M., R.E.D., and T.E.-C.]) (<http://www.anid.cl>), Las Salinas S.A. (M.S., R.O., X.B.-M., R.E.D., and T.E.-C.), and USM (M.S., X.B.-M., and R.E.D.) (<http://www.usm.cl>) grants. F.S.-S., D.J.-L., and E.R.B.M. were supported by the “Genomics and Proteomics Research on Bacterial Diversity” Culture Collection University of Gothenburg (CCUG) grant. The authors acknowledge the support of the National Genomics Infrastructure/Uppsala Genome Center and UPPMAX for providing assistance in massive parallel sequencing and computational infrastructure, funded by RFI/VR and Science for Life Laboratory, Sweden. This research was supported by the supercomputing infrastructure of the NLHPC (CCSS210001). The CCUG (www.ccug.se) is supported by the Department of Clinical Microbiology, Sahlgrenska University Hospital, and the Sahlgrenska Academy of the University of Gothenburg.

AUTHOR AFFILIATIONS

¹Molecular Microbiology and Environmental Biotechnology Laboratory, Department of Chemistry and Center of Biotechnology Daniel Alkalay Lowitt, Universidad Técnica Federico Santa María, Valparaíso, Chile

²Millennium Nucleus Bioproducts, Genomics and Environmental Microbiology (BioGEM), Valparaíso, Chile

³Laboratorio de Biología Celular y Ecofisiología Microbiana, Facultad de Ciencias Naturales y Exactas, Universidad de Playa Ancha, Valparaíso, Chile

⁴HUB Ambiental UPLA, Universidad de Playa Ancha, Valparaíso, Chile

⁵Department of Infectious Diseases, Institute of Biomedicine, University of Gothenburg, Sahlgrenska Academy, Gothenburg, Sweden

⁶Centre for Antibiotic Resistance Research (CARE), University of Gothenburg, Gothenburg, Sweden

⁷Culture Collection University of Gothenburg (CCUG), Department of Clinical Microbiology, Sahlgrenska University Hospital, Region Västra Götaland and Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

⁸Unit of Medical Technology and Biological Function, Department of Medical Technology and Diagnostics, RISE Research Institutes of Sweden, Mölndal, Sweden

AUTHOR ORCIDS

Francisco Salvà-Serra  <http://orcid.org/0000-0003-0173-560X>

Daniel Jaén-Luchoro  <http://orcid.org/0000-0002-5988-6227>

Edward R. B. Moore  <http://orcid.org/0000-0001-7693-924X>

Michael Seeger  <http://orcid.org/0000-0002-3925-1996>

FUNDING

Funder	Grant(s)	Author(s)
Agencia Nacional de Investigación y Desarrollo	ANID-Milenio-NCN2023_054	Ximena Báez-Matus Roberto E. Durán Roberto Orellana Teresa Esparza-Correa Michael Seeger
Fondo Nacional de Desarrollo Científico y Tecnológico	Fondecyt 1250809	Ximena Báez-Matus Roberto E. Durán Teresa Esparza-Correa Michael Seeger
Fondo Nacional de Desarrollo Científico y Tecnológico	Fondecyt 1200756	Ximena Báez-Matus Roberto E. Durán Teresa Esparza-Correa Michael Seeger
LAS SALINAS S.A.	Bioremediation Studies	Ximena Báez-Matus Roberto E. Durán Roberto Orellana Teresa Esparza-Correa Michael Seeger
CCUG	Genomics and Proteomics Research on Bacterial Diversity	Francisco Salvà-Serra Daniel Jaén-Luchoro Edward R. B. Moore

AUTHOR CONTRIBUTIONS

Ximena Báez-Matus, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Validation, Visualization, Writing – original draft, Writing – review and editing | Roberto E. Durán, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Writing – original draft, Writing – review and editing | Roberto Orellana, Conceptualization, Investigation, Methodology, Supervision, Writing – review and editing | Andrés Cumsille, Investigation, Methodology, Writing – review and editing | Ester G. Rivera, Investigation, Methodology | Teresa Esparza-Correa, Investigation, Methodology | Francisco Salvà-Serra, Data curation, Investigation, Methodology, Validation, Writing – review and editing |

Daniel Jaén-Luchoro, Data curation, Investigation, Methodology, Validation, Writing – review and editing | Edward R. B. Moore, Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – review and editing | Michael Seeger, Conceptualization, Formal analysis, Funding acquisition, Investigation, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

The complete genome sequence of *Rhodococcus erythropolis* ICBD2 has been deposited in DDBJ/ENA/GenBank under accession number [JBSDZJ000000000](https://doi.org/10.1093/jbmab/bzab000). The genome sequence described in this paper is the first version, BioProject [PRJNA1309126](https://doi.org/10.1093/bioinformatics/btq000). The raw sequencing reads are deposited and publicly available at the Sequence Read Archive, under accession numbers [SRR35078506](https://doi.org/10.1093/bioinformatics/btq000) (MiniON) and [SRR35078507](https://doi.org/10.1093/bioinformatics/btq000) (Illumina).

REFERENCES

- de Carvalho CCCR, da Fonseca MMR. 2005. The remarkable *Rhodococcus erythropolis*. *Appl Microbiol Biotechnol* 67:715–726. <https://doi.org/10.1007/s00253-005-1932-3>
- Thompson D, Cognat V, Goodfellow M, Koechler S, Heintz D, Carapito C, Van Dorsselaer A, Mahmoud H, Sangal V, Ismail W. 2020. Phylogenomic classification and biosynthetic potential of the fossil fuel-biodesulfurizing *Rhodococcus* strain IGT58. *Front Microbiol* 11:1417. <https://doi.org/10.3389/fmicb.2020.01417>
- Undabarrena A, Salvà-Serra F, Jaén-Luchoro D, Castro-Nallar E, Méndez KN, Valencia R, Ugalde JA, Moore ERB, Seeger M, Cámara B. 2018. Complete genome sequence of the marine *Rhodococcus* sp. H-CA8f isolated from Comau Fjord in Northern Patagonia, Chile. *Mar Genomics* 40:13–17. <https://doi.org/10.1016/j.margen.2018.01.004>
- Seeger M, Pieper D. 2010. Genetics of biphenyl biodegradation and co-metabolism of PCBs, p 1180–1199. *In* Timmis KN (ed), *Handbook of hydrocarbon and lipid microbiology*. Springer, Berlin, Heidelberg.
- Ponce BL, Latorre VK, González M, Seeger M. 2011. Antioxidant compounds improved PCB-degradation by *Burkholderia xenovorans* strain LB400. *Enzyme Microb Technol* 49:509–516. <https://doi.org/10.1016/j.enzymictec.2011.04.021>
- Méndez V, Sepúlveda M, Izquierdo-Fiallo K, Macaya CC, Esparza T, Báez-Matus X, Durán RE, Levicán G, Seeger M. 2025. Surfing in the storm: how *Paraburkholderia xenovorans* thrives under stress during biodegradation of toxic aromatic compounds and other stressors. *FEMS Microbiol Rev* 49:fua021. <https://doi.org/10.1093/femsre/fua021>
- Durán RE, Méndez V, Rodríguez-Castro L, Barra-Sanhueza B, Salvà-Serra F, Moore ERB, Castro-Nallar E, Seeger M. 2019. Genomic and physiological traits of the marine bacterium *Alcaligenes aquatilis* QD168 isolated from Quintero Bay, Central Chile, reveal a robust adaptive response to environmental stressors. *Front Microbiol* 10:528. <https://doi.org/10.3389/fmicb.2019.00528>
- Salvà Serra F, Salvà-Serra F, Svensson-Stadler L, Busquets A, Jaén-Luchoro D, Karlsson R, R. B. Moore E, Gomila M. 2018. A protocol for extraction and purification of high-quality and quantity bacterial DNA applicable for genome sequencing: a modified version of the Marmur procedure. *Protoc exch*. <https://doi.org/10.1038/protex.2018.084>
- Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. Available from: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Antipov D, Korobeynikov A, McLean JS, Pevzner PA. 2016. hybridSPAdes: an algorithm for hybrid assembly of short and long reads. *Bioinformatics* 32:1009–1015. <https://doi.org/10.1093/bioinformatics/btv688>
- Chklovski A, Parks DH, Woodcroft BJ, Tyson GW. 2023. CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning. *Nat Methods* 20:1203–1212. <https://doi.org/10.1038/s41592-023-01940-w>
- 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>
- Meier-Kolthoff JP, Göker M. 2019. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat Commun* 10:2182. <https://doi.org/10.1038/s41467-019-10210-3>
- Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK. 2016. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Anal Methods* 8:12–24. <https://doi.org/10.1039/C5AY02550H>