




# Draft Genome Sequence of *Rhodococcus erythropolis* VKPM Ac-1659, a Putative Oil-Degrading Strain Isolated from Polluted Soil in Siberia

A. A. Korzhenkov,<sup>a</sup> E. D. Bakhmutova,<sup>a</sup> A. O. Izotova,<sup>a</sup> A. A. Bavtushnyi,<sup>b</sup> K. V. Sidoruk,<sup>b</sup> E. V. Patrusheva,<sup>a,b</sup> M. V. Patrushev,<sup>a</sup>  S. V. Toshchakov<sup>a</sup>

<sup>a</sup>National Research Center "Kurchatov Institute," Moscow, Russia

<sup>b</sup>National Research Center "Kurchatov Institute"—GOSNIIGENETIKA, Moscow, Russia

**ABSTRACT** The draft genome sequence of *Rhodococcus erythropolis* VKPM Ac-1659, a putative oil-degrading strain, is reported. This genome sequence may provide better insights into the diversity and evolution of the genes responsible for hydrocarbon degradation in soil microorganisms.

Representatives of the species *Rhodococcus erythropolis*, first described by Gray and Thornton in 1928 as aromatic compound decomposers (1), are widely used in bioremediation technologies as bioemulsifiers and hydrocarbon degraders (2). Here, we report the draft genome sequence of *R. erythropolis* VKPM Ac-1659, a putative oil-degrading strain isolated from oil-polluted soil collected about 753 km from the city of Novosibirsk, Russia, and deposited in the Russian National Collection of Industrial Microorganisms (VKPM; <https://vkpm.genetika.ru/>).

For DNA extraction, the strain was grown on meat-peptone medium (3) at 29°C for 24 h. DNA was isolated using the ammonium salt treatment technique, as described previously (4). A genomic library was prepared using the KAPA HyperPlus kit (Roche, Switzerland) according to the manufacturer's recommendations. Sequencing was performed on an Illumina MiSeq system using a 2 × 250-bp format and a MiSeq reagent kit; 1,745,674 read pairs were obtained, yielding 866.4 million nucleotides.

Default parameters were used for all software unless otherwise specified. Quality control of the raw reads was performed using the fastp v0.20.1 package (5). Read processing and genome assembly were conducted using the ZGA pipeline (6), with the following steps: (i) quality read trimming and adapter removal using BBDuk v38.75, (ii) merging of the overlapping paired reads using bbmerge v38.75 (7), and (iii) *de novo* assembly of the assembled treated reads using the SPAdes v3.13.1 assembler (8). The genome assembly consisted of 170 contigs with a total length of 6,612,324 bp and an  $N_{50}$  value of 664,755 bp. The final genome coverage was 74×. The G+C content of the assembly was 62.38%. Analysis of the assembly quality using CheckM (9) showed both a high level of completeness (99.94%) and a low level of predicted contamination (3.24%). Genome-based taxonomic classification of the sequenced strains was performed using the Type Strain Genome Server (TYGS) (10). The TYGS analysis identified strain VKPM Ac-1659 as *Rhodococcus erythropolis* with high confidence; the estimated digital DNA-DNA hybridization (dDDH) values were greater than 85% for both the *R. erythropolis* type strains, NBRC 15567 and JCM 3201. Genome annotation was performed during submission of the data to NCBI using the NCBI Prokaryotic Genome Annotation Pipeline (11).

Analysis of the genes responsible for oil degradation was carried out using NCBI BLAST+ (12) with known oil degradation genes of both marine (13) and soil (14) microorganisms as queries. Two gene clusters were detected, including genes for rubredoxins, alkane 1-monooxygenase, and TetR family transcriptional regulators, the last of which might be

**Citation** Korzhenkov AA, Bakhmutova ED, Izotova AO, Bavtushnyi AA, Sidoruk KV, Patrusheva EV, Patrushev MV, Toshchakov SV. 2021. Draft genome sequence of *Rhodococcus erythropolis* VKPM Ac-1659, a putative oil-degrading strain isolated from polluted soil in Siberia. *Microbiol Resour Announc* 10:e00535-21. <https://doi.org/10.1128/MRA.00535-21>.

**Editor** J. Cameron Thrash, University of Southern California

**Copyright** © 2021 Korzhenkov et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to S. V. Toshchakov, [stepan.toshchakov@gmail.com](mailto:stepan.toshchakov@gmail.com).

**Received** 22 May 2021

**Accepted** 18 June 2021

**Published** 22 July 2021

responsible for the first step of alkane utilization, as described previously (15). One cluster consisted of AC1659\_RS04735, AC1659\_RS04740, AC1659\_RS04745, and AC1659\_RS04750, and the other consisted of AC1659\_RS17110, AC1659\_RS17115, AC1659\_RS17120, and AC1659\_RS17130. An additional contribution to alkane metabolism is possibly made by two NAD(P)/FAD-dependent oxidoreductases (AC1659\_RS20515 and AC1659\_RS22365) showing a high level of similarity with their counterparts in *Acinetobacter baylyi* ADP1, known to metabolize a wide spectrum of substrates, including phenol, benzyl alcohol, benzaldehyde, benzoate, and their hydroxy derivatives (16). In turn, the genes for polycyclic aromatic hydrocarbon degradation were not identified.

**Data availability.** The genome sequence has been deposited at NCBI GenBank and is available under accession number [JAGVVT0000000001](https://www.ncbi.nlm.nih.gov/nuccore/JAGVVT0000000001). The raw sequencing reads are available in NCBI SRA under accession number [SRX10910539](https://www.ncbi.nlm.nih.gov/sra/SRX10910539).

## ACKNOWLEDGMENTS

Sequencing was supported by a grant from the Ministry of Science and Higher Education of the Russian Federation allocated to the Kurchatov Center for Genome Research (agreement number 075-15-2019-1659). The genome analysis was supported by the NRC Kurchatov Institute (order number 1056 from 2 July 2020). The cultivation and DNA extraction were supported by the Russian State assignment number AAAA-A20-120093090016-9 and were carried out using the equipment of the unique scientific facility of BRC VKPM.

## REFERENCES

- Gray PHH, Thornton HG. 1928. Soil bacteria that decompose certain aromatic compounds. Zentralbl Bakteriell Parasitenk Infektionskr Hyg Abt II 73:74–96.
- Pacwa-Płociniczak M, Płociniczak T, Iwan J, Żarska M, Chorążewski M, Dzida M, Piotrowska-Seget Z. 2016. Isolation of hydrocarbon-degrading and biosurfactant-producing bacteria and assessment their plant growth-promoting traits. *J Environ Manage* 168:175–184. <https://doi.org/10.1016/j.jenvman.2015.11.058>.
- Atlas RM. 2004. Handbook of microbiological media. CRC Press, Boca Raton, FL.
- Sidoruk KV, Levitin EI, Sviridov BV, Pikasova OV, Shustikova T. 2020. Extraction of DNA from a wide range of objects by treatment with ammonium salts. *Biotekhnologija (Mosk)* 36:98–106. <https://doi.org/10.21519/0234-2758-2020-36-6-98-106>.
- Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ pre-processor. *Bioinformatics* 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>.
- Korzhenkov AA. 2021. ZGA: a flexible pipeline for read processing, de novo assembly and annotation of prokaryotic genomes. *bioRxiv* <https://doi.org/10.1101/2021.04.27.441618>.
- Bushnell B, Rood J, Singer E. 2017. BBMerge—accurate paired shotgun read merging via overlap. *PLoS One* 12:e0185056. <https://doi.org/10.1371/journal.pone.0185056>.
- 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>.
- 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>.
- 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>.
- Tatusova T, DiCuccio M, Badretdin A, Chetverin 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>.
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. <https://doi.org/10.1186/1471-2105-10-421>.
- Schneiker S, Dos Santos VAPM, Bartels D, Bekel T, Brecht M, Buhmester J, Chernikova TN, Denaro R, Ferrer M, Gertler C, Goesmann A, Golyshina OV, Kaminski F, Khachane AN, Lang S, Linke B, McHardy AC, Meyer F, Nechitaylo T, Pühler A, Regenhart D, Rupp O, Sabirova JS, Selbitschka W, Yakimov MM, Timmis KN, Vorhölter F-J, Weidner S, Kaiser O, Golyshin PN. 2006. Genome sequence of the ubiquitous hydrocarbon-degrading marine bacterium *Alcanivorax borkumensis*. *Nat Biotechnol* 24:997–1004. <https://doi.org/10.1038/nbt1232>.
- Laczi K, Kis Á, Horváth B, Maróti G, Hegedüs B, Perei K, Rákhely G. 2015. Metabolic responses of *Rhodococcus erythropolis* PR4 grown on diesel oil and various hydrocarbons. *Appl Microbiol Biotechnol* 99:9745–9759. <https://doi.org/10.1007/s00253-015-6936-z>.
- Toshchakov SV, Korzhenkov AA, Chernikova TN, Ferrer M, Golyshina OV, Yakimov MM, Golyshin PN. 2017. The genome analysis of *Oleiphilus messinensis* ME102 (DSM 13489<sup>T</sup>) reveals backgrounds of its obligate alkane-devouring marine lifestyle. *Mar Genomics* 36:41–47. <https://doi.org/10.1016/j.margen.2017.07.005>.
- Fewson CA. 1985. Growth yields and respiratory efficiency of *Acinetobacter calcoaceticus*. *Microbiology* 131:865–872. <https://doi.org/10.1099/00221287-131-4-865>.