AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

# Complete Genome Sequence of *Rhodococcus qingshengii* VT6, a Promising Degrader of Persistent Pollutants and Putative Biosurfactant-Producing Strain

**Yanina Delegan,[a] Kirill Petrikov,[a] Ekaterina Frantsuzova,[a,b] Alexander Bogun,[c] Vasili Travkin,[d] Inna Solyanikova[a,b,d]**

[a]Federal Research Center "Pushchino Scientific Center for Biological Research of the Russian Academy of Sciences," Pushchino, Moscow Region, Russian Federation
[b]The Federal State Budget Educational Institution of Higher Education Pushchino State Institute of Natural Science, Pushchino, Moscow Region, Russian Federation
[c]State Research Center for Applied Microbiology and Biotechnology, Obolensk, Russian Federation
[d]Belgorod State University, Belgorod, Russian Federation

**ABSTRACT** The strain *Rhodococcus qingshengii* VT6 is a promising degrader of persistent pollutants and a putative biosurfactant producer. The genome of the strain was sequenced completely. It consists of a 6,457,868-bp chromosome and 4 plasmids (pLP1, 501,672 bp; pLP2, 188,969 bp; pCP1, 100,387 bp; and pCP2, 132,858 bp).

Rhodococci are known to be common representatives of soil microflora (1). They are known as metabolically versatile microorganisms with potential applications in bioremediation (2).

The *Rhodococcus qingshengii* strain VT6 (VKM Ac-2909D) was isolated from a forest soil sample (54°83′20″, 37°61′60″; Moscow Region, Russia Federation). The strain transforms hexadecane and trinitrotoluene (3) and produces surface-active compounds. Surface tension (ST) was measured by the du Noüy method (4) using the tensiometer K6 (Kruss, Germany) at the temperature of 25°C. Cultivation of the strain in liquid Evans (5) medium (reference solution, ST 77 $mN \cdot m^{-1}$) with hexadecane 2% (vol/vol) at 25°C for 5 days resulted in the reduction of ST to 36 $mN \cdot m^{-1}$, which may indicate the synthesis of surfactants.

For long-term storage, the strain was kept in glycerol (40%) at −70°C. For short-term maintenance, the strain was cultured on LB (6) agar plates at 27°C.

Genomic DNA was isolated from a fresh biomass of *Rhodococcus* VT6 grown on LB agar using a DNeasy kit (catalog [cat.] no 69506; Qiagen). The 16S rRNA gene sequencing was performed as described in reference 3. The analysis of the sequencing results showed that the strain is related to *Rhodococcus erythropolis* or *R. qingshengii*. The closest relative of the VT6 strain is *R. qingshengii* TG-1 (GenBank accession no. CP077417.1).

Sequencing was performed in Federal Research Center "Pushchino Scientific Center for Biological Research, RAS" using a MinION sequencer with the flow cell R9.4.1 (Oxford Nanopore Technologies). A library was prepared with a ligation kit (SQK-LSK109). Guppy 3.2.4 was used for base calling, which yielded a total of 1,306.8 Mb distributed in 130,800 reads with a Q of >10 ($N_{50}$ is 17,156 bp).

The same DNA sample was sequenced with an Illumina NovaSeq6000 instrument using an S2 reagent kit (cat. no. 20012861) in BioSpark (Troitsk, Russia). A paired-end library was prepared with the HyperPlus kit (Kapa Biosystems). We obtained 37,607,400 paired-end reads of <101 bp. Quality control was performed using FastQC (7) and NanoPack (8). The Illumina and Nanopore reads were used for hybrid assembly with SPAdes 3.15.0 (9). The Nanopore reads were assembled into 5 contigs using Flye 2.9 (10). Next, SPAdes contigs were combined into replicons using Snapgene 6.0 with Flye data as the reference. The Illumina reads were used to correct Nanopore errors using Bowtie 2 2.4.4 (11) and Pilon 1.24 (12). Default parameters were used for all software unless otherwise specified.

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



Address correspondence to Yanina Delegan, mewgia@ya.ru.

The authors declare no conflict of interest.

**Received** 7 December 2021
**Accepted** 21 January 2022
**Published** 10 February 2022

The *R. qingshengii* VT6 genome consists of a 6,457,868-bp chromosome (GC content, 62.41%) and 4 plasmids. The plasmids pLP1 (501,672 bp; GC content, 61.50%) and pLP2 (188,969 bp; GC content, 61.47%) are linear, pCP1 (100,387 bp; GC content, 61.88%) and pCP2 (132,858 bp; GC content, 63.24%) are circular. All plasmids also were visualized using pulsed-field electrophoresis. Chromosome and plasmid circularization were specified by end overlapping and using Tablet 1.21.02.08 (13). The ends of linear plasmids were verified by amplification (see supplemental data online at FigShare [10.6084/m9.figshare.18758252]) and subsequent sequencing.

To identify the strain to species, we used average nucleotide identity (ANI) value (https://www.ezbiocloud.net/tools/ani [14]) and digital DNA-DNA hybridization (DDH) (https://ggdc.dsmz.de/ggdc.php [15]). The ANI value and DDH with the type strain *R. erythropolis* NBRC15567 are 95.38% and 75.60%, respectively, and with the type strain *R. qingshengii* JCM15477 are 98.45% and 77.30%, respectively. So, we identify the strain VT6 as *R. qingshengii*.

The genome of strain VT6 was annotated using NCBI PGAP 4.6 (16) and Prokka 1.14.6 (17). The strain bears a number of catabolic genes for alkane degradation. We found a set of trehalose biosynthesis genes. Trehalose is one of the main components of glycolipid biosurfactants. The VT6 strain is a potential producer of biosurfactants.

**Data availability.** This genome project has been deposited at GenBank under BioSample SAMN23500510, BioProject PRJNA784759, GenBank accession numbers CP088906 to CP088910, and SRA accession numbers SRX13270910 for Illumina and SRX13270985 for Oxford Nanopore data.

## ACKNOWLEDGMENT

This work was financially supported by Ministry of Science and Higher Education of the Russian Federation (agreement no. 075-15-2021-973, government contract no. 13.2251.21.0039).

## REFERENCES

1. Larkin MJ, Kulakov LA, Allen CC. 2006. Biodegradation by members of the genus *Rhodococcus*: biochemistry, physiology, and genetic adaptation. Adv Appl Microbiol 59:1–29. https://doi.org/10.1016/S0065-2164(06)59001-X.
2. Kim D, Choi KY, Yoo M, Zylstra GJ, Kim E. 2018. Biotechnological potential of *Rhodococcus* biodegradative pathways. J Microbiol Biotechnol 28:1037–1051. https://doi.org/10.4014/jmb.1712.12017.
3. Travkin V, Morudullaev D, Artemyeva I, Suzina N, Solyanikova I. 2021. Soil bacteria as a basis for sustainable development of the environment. E3S Web Conf 247:01051–01055. https://doi.org/10.1051/e3sconf/202124701051.
4. Petrikov K, Delegan Y, Surin A, Ponamoreva O, Puntus I, Filonov A, Boronin A. 2013. Glycolipids of *Pseudomonas* and *Rhodococcus* oil-degrading bacteria used in bioremediation preparations: formation and structure. Process Biochem 48:931–935. https://doi.org/10.1016/j.procbio.2013.04.008.
5. Delegan YA, Valentovich LN, Shafieva SM, Ganbarov KG, Filonov AE, Vainstein MB. 2019. Characterization and genomic analysis of highly efficient thermotolerant oil-degrading bacterium *Gordonia* sp. 1D. Folia Microbiol (Praha) 64:41–48. https://doi.org/10.1007/s12223-018-0623-2.
6. Bertani G. 1951. Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. J Bacteriol 62:293–300. https://doi.org/10.1128/jb.62.3.293-300.1951.
7. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc.
8. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. Bioinformatics 34:2666–2669. https://doi.org/10.1093/bioinformatics/bty149.
9. Bankevich A, Nurk S, Antipov D, Gurevich A, Dvorkin M, Kulikov AS, Lesin V, Nikolenko S, Pham S, Prjibelski A, Pyshkin A, Sirotkin A, 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.
10. 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.
11. Langmead B, Salzberg S. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923.
12. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9:e112963. https://doi.org/10.1371/journal.pone.0112963.
13. Milne I, Stephen G, Bayer M, Cock PJA, Pritchard L, Cardle L, Shaw PD, Marshall D. 2013. Using Tablet for visual exploration of second-generation sequencing data. Brief Bioinform 14:193–202. https://doi.org/10.1093/bib/bbs012.
14. Yoon SH, Ha SM, Lim JM, Kwon SJ, Chun J. 2017. A large-scale evaluation of algorithms to calculate average nucleotide identity. Antonie Van Leeuwenhoek 110:1281–1286. https://doi.org/10.1007/s10482-017-0844-4.
15. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 14:60. https://doi.org/10.1186/1471-2105-14-60.
16. 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.
17. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.