



Genome Sequence of *Rhodococcus erythropolis* Type Strain JCM 3201

Keitaro Yoshida,^a Wataru Kitagawa,^{a,b,c} Koji Ishiya,^b Yasuo Mitani,^a Nobutaka Nakashima,^{a,b} Sachiyo Aburatani,^{b,d} Tomohiro Tamura^{a,b,c}

^aBioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Sapporo, Japan

^bComputational Bio Big-Data Open Innovation Laboratory (CBBDOIL), AIST, Tokyo, Japan

^cGraduate School of Agriculture, Hokkaido University, Sapporo, Japan

^dBiotechnology Research Institute for Drug Discovery, AIST, Tokyo, Japan

ABSTRACT *Rhodococcus erythropolis* JCM 3201 can express several recombinant proteins that are difficult to express in *Escherichia coli*. It is used as one of the hosts for protein expression and bioconversion. Here, we report the draft genome sequence of *R. erythropolis* JCM 3201.

Rhodococci remarkably degrade various xenobiotics (1–3). The type strain of *Rhodococcus erythropolis*, JCM 3201, was originally isolated from soil as a degrader of aromatics (4). A lysozyme-sensitive mutant of JCM 3201, strain L88, has been used as an expression host of recombinant proteins that are difficult to express in *Escherichia coli* (5–9). JCM 3201 is also useful to the host for bioconversion of chemicals such as vitamin D₃ to hydroxyvitamin D₃ (10–12). To optimize these processes through metabolic pathway engineering, its genomic information will be essential. The draft genome of JCM 3201 was previously deposited with 67 contigs (GenBank accession number [BCRM00000000](https://doi.org/10.1128/MRA.01730-18)). However, this sequence lacks information about the numbers and structural type of its chromosomes and plasmids.

To improve the genome assembly level, we resequenced the JCM 3201 genome. We cultured the strain in LB broth containing 1% glycine, lysed cells with 2 mg ml⁻¹ lysozyme, and extracted the genomic DNA by the phenol-chloroform method (13). We prepared three libraries with the TruSeq DNA PCR-free library prep kit (Illumina), the Nextera mate pair sample preparation kit (Illumina), and the DNA template prep kit version 1.0 (PacBio). The genome was sequenced using (i) an Illumina HiSeq 2500 platform with 350-bp paired-end and 8-kbp mate pair libraries and (ii) a PacBio RS II platform. The PacBio raw reads were filtered with SMRT Analysis version 2.3.0 with the following parameters: minimum subread length, 500; minimum polymerase read quality, 0.80; and minimum polymerase read length, 100. We obtained 98,981 subreads and assembled them with the Hierarchical Genome Assembly Process (HGAP) version 2 (14), resulting in a 6.33-Mbp circular sequence with a mean coverage of 84-fold. The paired-end and mate pair reads were filtered with Trimmomatic version 0.38 with the parameters SLIDINGWINDOW:20:20 and MINLEN:50 (15), resulting in 14,661,238 and 9,431,622 reads, respectively. We assembled these reads using Velvet version 1.2.08 (16) and obtained 52 contigs. Among these contigs, 12 did not match the circular PacBio assembly, which showed the highest similarities to plasmid sequences of other *R. erythropolis* strains (17). To confirm the possibility that these contigs were derived from plasmid sequences, we examined their connectivity by PCR and finally obtained another 85-kb circular sequence and a 241-kb linear sequence. Based on their sizes, we predicted that these two sequences were plasmids. One end of the linear sequence showed similarity to end sequences of linear plasmids in rhodococci. However, the

Citation Yoshida K, Kitagawa W, Ishiya K, Mitani Y, Nakashima N, Aburatani S, Tamura T. 2019. Genome sequence of *Rhodococcus erythropolis* type strain JCM 3201. *Microbiol Resour Announc* 8:e01730-18. <https://doi.org/10.1128/MRA.01730-18>.

Editor Jason E. Stajich, University of California, Riverside

Copyright © 2019 Yoshida 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 Tomohiro Tamura, t-tamura@aist.go.jp.

Received 3 January 2019

Accepted 27 February 2019

Published 4 April 2019

other end did not, suggesting that this linear sequence might not be completed yet. To polish the assembled sequences, we mapped the filtered paired-end reads to the initial assembly using the Burrows-Wheeler Aligner MEM algorithm (BWA-MEM) version 0.7.12 with a seed length of 19 nucleotides (18) and corrected errors using the Genome Analysis Toolkit (GATK) version 4.0.6.0 of variant filtration with default parameters (19). We annotated the genes using DFAST (20).

The draft genome sequence of JCM 3201 contained two circular sequences and a linear sequence with a length of 6,326,569 bp (62.4% G+C content), 84,587 bp (62.5% G+C content), and 240,958 bp (61.2% G+C content), respectively. It contains 6,152 putative coding DNA sequences (CDSs), 12 rRNAs, and 53 tRNAs.

Data availability. The DDBJ/EMBL/GenBank sequence accession numbers for this project are [BHXB01000001](#) to [BHXB01000003](#). The SRA accession numbers are [DRX143934](#) to [DRX143936](#).

ACKNOWLEDGMENT

This study was supported by an internal grant from AIST.

REFERENCES

- Kitagawa W, Miyauchi K, Masai E, Fukuda M. 2001. Cloning and characterization of benzoate catabolic genes in the Gram-positive polychlorinated biphenyl degrader *Rhodococcus* sp. strain RHA1. *J Bacteriol* 183: 6598–6606. <https://doi.org/10.1128/JB.183.22.6598-6606.2001>.
- Kitagawa W, Kimura N, Kamagata Y. 2004. A novel *p*-nitrophenol degradation gene cluster from a Gram-positive bacterium, *Rhodococcus opacus* SAO101. *J Bacteriol* 186:4894–4902. <https://doi.org/10.1128/JB.186.15.4894-4902.2004>.
- Kitagawa W, Suzuki A, Hoaki T, Masai E, Fukuda M. 2001. Multiplicity of aromatic ring hydroxylation dioxygenase genes in a strong PCB degrader, *Rhodococcus* sp. strain RHA1 demonstrated by denaturing gradient gel electrophoresis. *Biosci Biotechnol Biochem* 65:1907–1911. <https://doi.org/10.1271/bbb.65.1907>.
- Goodfellow M, Alderson G. 1977. The actinomycete-genus *Rhodococcus*: a home for the “*rhodochrous*” complex. *J Gen Microbiol* 100:99–122. <https://doi.org/10.1099/00221287-100-1-99>.
- Nakashima N, Mitani Y, Tamura T. 2005. Actinomycetes as host cells for production of recombinant proteins. *Microb Cell Fact* 4:7. <https://doi.org/10.1186/1475-2859-4-7>.
- Mitani Y, Nakashima N, Sallam KI, Toriyabe T, Kondo K, Tamura T. 2006. Advances in the development of genetic tools for the genus *Rhodococcus*. *Actinomycetologica* 20:55–61. <https://doi.org/10.3209/saj.20.55>.
- Nakashima N, Tamura T. 2004. A novel system for expressing recombinant proteins over a wide temperature range from 4 to 35°C. *Biotechnol Bioeng* 86:136–148. <https://doi.org/10.1002/bit.20024>.
- Mitani Y, Meng X, Kamagata Y, Tamura T. 2005. Characterization of LtsA from *Rhodococcus erythropolis*, an enzyme with glutamine amidotransferase activity. *J Bacteriol* 187:2582–2591. <https://doi.org/10.1128/JB.187.8.2582-2591.2005>.
- Vallecillo AJ, Parada C, Morales P, Espitia C. 2017. *Rhodococcus erythropolis* as a host for expression, secretion and glycosylation of *Mycobacterium tuberculosis* proteins. *Microb Cell Fact* 16:12. <https://doi.org/10.1186/s12934-017-0628-6>.
- Fujii Y, Kabumoto H, Nishimura K, Fujii T, Yanai S, Takeda K, Tamura N, Arisawa A, Tamura T. 2009. Purification, characterization, and directed evolution study of a vitamin D₃ hydroxylase from *Pseudonocardia autotrophica*. *Biochem Biophys Res Commun* 385:170–175. <https://doi.org/10.1016/j.bbrc.2009.05.033>.
- Imoto N, Nishioka T, Tamura T. 2011. Permeabilization induced by lipid
- Il-targeting lantibiotic nisin and its effect on the bioconversion of vitamin D₃ to 25-hydroxyvitamin D₃ by *Rhodococcus erythropolis*. *Biochem Biophys Res Commun* 405:393–398. <https://doi.org/10.1016/j.bbrc.2011.01.038>.
- Yasutake Y, Nishioka T, Imoto N, Tamura T. 2013. A single mutation at the ferredoxin binding site of P450 Vdh enables efficient biocatalytic production of 25-hydroxyvitamin D₃. *ChemBioChem* 14:2284–2291. <https://doi.org/10.1002/cbic.201300386>.
- Sambrook J, Russell RW. 2001. *Molecular cloning: a laboratory manual*, 3rd ed, vol 1. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <https://doi.org/10.1038/nmeth.2474>.
- 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>.
- Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res* 18:821–829. <https://doi.org/10.1101/gr.074492.107>.
- Rückert C, Birmes FS, Müller C, Niewerth H, Winkler A, Fetzner S, Kalinowski J. 2015. Complete genome sequence of *Rhodococcus erythropolis* BG43 (DSM 46869), a degrader of *Pseudomonas aeruginosa* quorum sensing signal molecules. *J Biotechnol* 211:99–100. <https://doi.org/10.1016/j.jbiotec.2015.07.014>.
- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>.
- Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, Banks E, Garimella KV, Altshuler D, Gabriel S, DePristo MA. 2013. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* 43:11.10.1–11.10.33. <https://doi.org/10.1002/0471250953.bi1110s43>.
- Tanizawa Y, Fujisawa T, Nakamura Y. 2018. DFAST: a flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics* 34:1037–1039. <https://doi.org/10.1093/bioinformatics/btx713>.