

Draft Genome Sequence of *Rhodococcus qingshengii* Strain BKS 20-40

Monu Bala,^a Shailesh Kumar,^b Gajendra Pal Singh Raghava,^b Shanmugam Mayilraj^a

Microbial Type Culture Collection and Gene Bank (MTCC), Chandigarh, India^a; Bioinformatics Centre, CSIR-Institute of Microbial Technology, Chandigarh, India^b

M.B. and S.K. contributed equally to this article.

This is IMTECH communication number 03/2013.

We report the 5.8-Mb genome sequence of *Rhodococcus qingshengii* strain BKS 20-40, isolated from a palm tree rhizosphere soil sample from Bhitarkanika National Park, Odisha, India. The strain is capable of degrading cholesterol moiety. The draft genome of strain BKS 20-40 consists of 6,601,618 bp, with 62.4% G+C content.

Received 25 February 2013 Accepted 28 February 2013 Published 28 March 2013

Citation Bala M, Kumar S, Raghava GPS, Mayilraj S. 2013. Draft genome sequence of *Rhodococcus qingshengii* strain BKS 20-40. *Genome Announc.* 1(2):e00128-13. doi:10.1128/genomeA.00128-13.

Copyright © 2013 Bala et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported License](http://creativecommons.org/licenses/by/3.0/).

Address correspondence to Shanmugam Mayilraj, mayil@imtech.res.in.

The genus *Rhodococcus* was first proposed by Zopf (1891) and amended by Tsukamura (1974) (1) and Goodfellow and Alderson (1977) (2). *Rhodococcus qingshengii* was first characterized by Jing-Liang Xu et al. (3). *R. qingshengii* strain BKS 20-40 is a Gram-positive, aerobic, nonmotile, mesophilic strain, isolated from a palm tree rhizosphere soil sample collected from the Bhitarkanika Mangrove Reserve Forest, Odisha, India. During screening, we found that this strain is capable of producing cholesterol oxidase. The genome of *R. qingshengii* BKS 20-40 was sequenced using the Illumina-HiSeq 1000 paired-end technology, which produced a total of 28,206,978 paired-end reads (insert size of 350 bp) of 101 bp. We have used the next-generation sequencing quality-control NGS QC toolkit v2.3 (4) to filter the data for high-quality (HQ) (cutoff read length for HQ = 70%; cutoff quality score = 20) vector- and adaptor-free reads for genome assembly. A total of 26,791,554 high-quality vector-filtered reads (~541.2× coverage) were used for assembly with Velvet 1.2.08 (at a hash length of 55) (5). The final assembly contains 104 contigs of total size 6,601,618 bp, with an N₅₀ contig length of 175.3 kb; the largest contigs assembled measure 483.3 kb.

The draft genome (104 contigs) comprising 6,601,618 bp was annotated with the help of the Rapid Annotations using Subsystems Technology (RAST) (6) and RNAMmer 1.2 (7) servers. A total of 6,409 predicted coding regions (CDSs), 3 rRNAs, and 52 tRNAs were predicted.

RAST annotation indicates that strain *Rhodococcus jostii* RHA1 (score 511), *Rhodococcus erythropolis* PR4 (score 494), and *Rhodococcus erythropolis* SK121 are the closest neighbors of strain BKS 20-40. RAST server annotation also indicates that strain BKS 20-40 has the genes coding for glycolysis, gluconeogenesis, and the tricarboxylic acid (TCA) cycle. Strain BKS 20-40 also has the genes encoding the acetyl-coenzyme A (CoA) fermentation to butyrate, including enoyl-CoA hydratase (EC 4.2.1.17), acetyl-CoA acetyltransferase (EC 2.3.1.9), butyryl-CoA dehydrogenase (EC 1.3.99.2), 3-hydroxybutyryl-CoA epimerase (EC 5.1.2.3), 3-hydroxybutyryl-CoA dehydrogenase (EC 1.1.1.157), 3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35), and electron transfer flavoprotein (both al-

pha and beta subunits). In the annotation, we have also found the genes encoding 3-ketoacyl-CoA thiolase (EC 2.3.1.16), long-chain-fatty-acid-CoA ligase (EC 6.2.1.3), and aldehyde dehydrogenase (EC 1.2.1.3).

Nucleotide sequence accession numbers. The draft genome sequence of *R. qingshengii* BKS 20-40 has been included in the GenBank Whole-Genome Shotgun (WGS) database under the accession no. [AODN000000000](http://www.ncbi.nlm.nih.gov/nuccore/AODN000000000). The version described in this paper is the first version, accession no. [AODN010000000](http://www.ncbi.nlm.nih.gov/nuccore/AODN010000000).

ACKNOWLEDGMENTS

This work was funded by IMTECH-CSIR. M.B. is supported by a University Grant Commission (UGC) fellowship, and S.K. is supported by a research fellowship from the Council of Scientific and Industrial Research.

We thank the C-CAMP (<http://www.ccamp.res.in/>) next-generation genomics facility for help in obtaining the genome sequence.

Genome assembly and annotation data can be downloaded from our genomics web portal at <http://crdd.osdd.net/raghava/genomesrs/>.

REFERENCES:

1. Tsukamura M. 1974. A further numerical taxonomic study of the rhodochrous group. *Jpn. J. Microbiol.* 18:37–44.
2. Goodfellow M, Alderson G. 1977. The actinomycete-genus *Rhodococcus*: a home for the “rhodochrous” complex. *J. Gen. Microbiol.* 100:99–122.
3. Xu JL, He J, Wang ZC, Wang K, Li WJ, Tang SK, Li SP. 2007. *Rhodococcus qingshengii* sp. nov., a carbendazim-degrading bacterium. *Int. J. Syst. Evol. Microbiol.* 57:2754–2757.
4. Patel RK, Jain M. 2012. NGS QC toolkit: a toolkit for quality control of next generation sequencing data. *PLoS One* 7:e30619. doi:10.1371/journal.pone.0030619.
5. Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res.* 18:821–829.
6. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9:75.
7. Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. 2007. RNAMmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35:3100–3108.