JB Journals.ASM.org

GENOME ANNOUNCEMENT

Draft Genome Sequence of the Nitrophenol-Degrading Actinomycete *Rhodococcus imtechensis* RKJ300

Surendra Vikram, Shailesh Kumar, Srikrishna Subramanian, and Gajendra Pal Singh Raghava

Council of Scientific and Industrial Research—Institute of Microbial Technology, Sector 39-A, Chandigarh, India

We report the 8.231-Mb genome sequence of *Rhodococcus imtechensis* RKJ300, isolated from pesticide-contaminated soil in Punjab, India. The genome sequence of the strain RKJ300 will be helpful in exploring the molecular pathways involved in the degradation of nitrophenols.

Rhodococcus imtechensis RKJ300 (MTCC 7085) was isolated from pesticide-contaminated soil in Punjab, India, by enrichment on minimal medium containing 4-nitrophenol (3). Strain RKJ300 is capable of utilizing 4-nitrophenol, 2-chloro-4-nitrophenol, and 2, 4-dinitrophenol as sole sources of carbon and energy (2). The strain uses both oxidative and reductive catabolic mechanisms for initial transformation of these compounds. Due to its role in biodegradation, the strain RKJ300 was selected for whole-genome sequencing.

The genome of *Rhodococcus imtechensis* RKJ300 was sequenced using Illumina-HiSeq 1000 paired end technology, which produced a total of 30,464,548 paired-end reads of 101 bp. We used the NGS QC toolkit v2.2.1 (6) to filter the data for high-quality (cutoff read length for HQ = 70%, cutoff quality score = 20), vector- and adaptor-free reads for genome assembly. A total of 22,040,838 high-quality, vector-filtered reads (~267× coverage) was used for assembly with SOAPdenovo v1.05 (5) and Gap Closer (for closing the gaps after scaffolding with SOAPdenovo v1.05). A total of 175 scaffolds of 8,231,486 nucleotides with an *N*50 of 135.8 kb were produced after the assembly. Due to the presence of 146 gaps (*N*s), 175 scaffolds were split into 178 contigs (over 200 bp) of 8.231 Mb with an *N*50 of 135.8 kb. The data were submitted to GenBank.

The genome of strain RKJ300 has a G+C content of 67.22%, and annotation using the RAST (rapid annotation using subsystem technology) system (1) and RNAmmer 1.2 (4) servers revealed 8,059 predicted coding regions (CDSs), 49 tRNAs, and 5 rRNA genes. Comparison of genome sequence available at the RAST server shows that *Rhodococcus jostii* RHA1 (score, 548), *Rhodococcus opacus* B4 (score, 421), and *Rhodococcus erythropolis* PR4 (score, 417) are the closest neighbors of RKJ300.

RKJ300 has the genes for cysteine desulfurase (EC 2.8.1.7), D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95), glycerate kinase (EC 2.7.1.31), 3-ketoacyl coenzyme A (CoA) thiolase (EC 2.3.1.16), butyryl CoA dehydrogenase (EC 1.3.99.2), enoyl-CoA hydratase (EC 4.2.1.17), butyryl-CoA dehydrogenase (EC 1.3.99.2), enoyl-CoA hydratase (EC 4.2.1.17), aldehyde dehydrogenase (EC 1.2.1.3), enoyl-CoA hydratase (EC 4.2.1.17), long-chain-fatty-acid-CoA ligase (EC 6.2.1.3),and enoyl-CoA hydratase (EC 4.2.1.17,) which have also been reported to occur in strain RHA1 (5a). In contrast to *Rhodococcus jostii* RHA1, genes for nitrilotriacetate monooxygenase component B (EC 1.14.13.-), 3-oxoadipate CoA-transferase subunit A (EC 2.8.3.6), 2-hydroxy-6-oxo-6-phenylhexa-2, 4-dienoate hydrolase (EC 3.7.1.-), 2-polyprenylphenol hydroxylase, and related flavodoxin oxidoreductases and 4-oxalocrotonate tautomerase (EC 5.3.2.-) are present only in strain RKJ300. We also manually found the genes of benzoylformate decarboxylase (EC 4.1.1.7), 4-hydroxyphenylacetate 3-monooxygenase (EC 1.14.13.3), 4-nitrophenylphosphatase, benzoate 1,2-dioxygenase alpha subunit (EC 1.14.12.10), benzoate 1,2-dioxygenase beta subunit (EC 1.14.12.10), *S*-nitrosomycothiol reductase MscR and *para*-nitrobenzyl esterase (EC 3.1.1.-), which are involved in the biodegradation of aromatic compounds.

Genome assembly and annotation data files can be downloaded from the web portal at http://crdd.osdd.net/raghava/genomesrs/.

Nucleotide sequence accession numbers. This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession AJJH00000000. The version described in this paper is the first version, AJJH01000000.

ACKNOWLEDGMENTS

This work was funded by IMTECH-CSIR. S.K. and S.V. are supported by a research fellowship from the Council of Scientific and Industrial Research.

We thank P. Anil Kumar, IMTECH, Chandigarh, India, for providing help in the genomic DNA extraction. We also thank Centre for Cellular and Molecular Platforms (C-CAMP), a Department of Biotechnology (Government of India) initiative, for providing the high quality Illumina-HiSeq 1000 data.

REFERENCES

1. **Aziz RK, et al.** 2008. The RAST server: rapid annotations using subsystems technology. BMC Genomics **9**:75.
2. **Ghosh A, et al.** 2010. Degradation of 4-nitrophenol, 2-chloro-4-nitrophenol, and 2, 4-dinitrophenol by *Rhodococcus imtechensis* strain RKJ300. Environ. Sci. Technol. **44**:1069–1077.
3. **Ghosh A, Paul D, Prakash D, Mayilraj S, Jain RK.** 2006. *Rhodococcus imtechensis* sp. nov., a nitrophenol-degrading actinomycete. Int. J. Syst. Evol. Microbiol. **56**:1965–1969.
4. **Lagesen K, et al.** 2007. RNAmmer: consistent annotation of rRNA genes in genomic sequences. Nucleic Acids Res. **35**:3100–3108.
5. **Li R, et al.** 2010. De novo assembly of human genomes with massively parallel short read sequencing. Genome Res. **20**:265–272.

5a. **McLeod MP, et al.** 2006. The complete genome of Rhodococcus sp. RHA1 provides insights into a catabolic powerhouse. Proc. Natl. Acad. Sci. U. S. A. **103**:15582–15587.

6. **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.

Received 2 April 2012 **Accepted** 11 April 2012

Address correspondence to Gajendra Pal Singh Raghava, raghava@imtech.res.in.

S.V. and S.K. contributed equally to this article.

doi:10.1128/JB.00532-12