JB Journals.ASM.org

GENOME ANNOUNCEMENT

Draft Genome Sequence and Comparative Analysis of the Superb Aromatic-Hydrocarbon Degrader *Rhodococcus* sp. Strain DK17

Miyoun Yoo,[a,b] Dockyu Kim,[c] Ki Young Choi,[a]* Jong-Chan Chae,[d] Gerben J. Zylstra,[e] and Eungbin Kim[a,b]

Department of Systems Biology, Yonsei University, Seoul, South Korea[a]; Yonsei–Paris-Sud Project, Yonsei International Campus, Incheon, South Korea[b]; Division of Life Sciences, Korea Polar Research Institute, Incheon, South Korea[c]; Division of Biotechnology, Chonbuk National University, Iksan, South Korea[d]; and Department of Biochemistry and Microbiology, School of Environmental and Biological Sciences, Rutgers University, New Brunswick, New Jersey, USA[e]

***Rhodococcus* sp. strain DK17 is capable of utilizing various derivatives of benzene and bicyclics containing both aromatic and alicyclic moieties as sole carbon and energy sources. Here, we present the 9,107,362-bp draft genome sequence of DK17 and its genomic analysis in comparison with other members of the genus *Rhodococcus*.**

Members of the genus *Rhodococcus* demonstrate a remarkable ability to utilize a wide variety of natural organic and xenobiotic compounds, including aliphatic, aromatic, and alicyclic hydrocarbons (7). *Rhodococcus* sp. strain DK17 was isolated from a crude oil-contaminated site in Yeochon, South Korea, for the ability to grow on *o*-xylene as the sole source of carbon and energy and was later found to also have the ability to grow on benzene, toluene, ethylbenzene, isopropylbenzene, *n*-propyl- to *n*-hexylbenzenes, phenol, indan, tetralin, and phthalates (3, 5, 6). Strain DK17 possesses three linear megaplasmids (380-kb pDK1, 330-kb pDK2, and 750-kb pDK3). The genes for alkylbenzene degradation are present on pDK2, while the gene clusters for phthalate degradation are duplicated and found on both pDK2 and pDK3 (2, 4).

In an effort to fully understand the metabolic versatility of DK17, its genome was sequenced using Roche/454 GS FLX (577,562 reads) and Illumina GAIIx (32,685,616 reads) systems and assembled with Newbler assembler version 2.3 software (454 Life Sciences) and CLC Genomics Workbench version 4.5.1 (Illumina GAIIx) with totals of 9,107,362 bases in paired-end reads (totaling ~9 Mb; ~567-fold coverage of the genome). Optical mapping was performed to confirm the assembly output. The annotation was done by merging the results obtained from the RAST (Rapid Annotation using Subsystem Technology) server (1) and COG (Cluster of Orthologous Groups) database (8). Analysis of the unclosed draft genome sequence of strain DK17, consisting of 135 contigs in 27 scaffolds, showed that the G+C content is 67.15% and that the sequence contains 9,266 open reading frames (ORFs). Additionally, at least 48 tRNA genes and 4 rRNA operons were predicted in the draft genome. The ORF annotation by COG and RAST results in classification into 18 COG categories (J, K, L, D, O, M, N, P, T, C, G, E, F, H, I, Q, R, and S) with a total of 6,124 genes.

Comparison to the genome sequences of other *Rhodococcus* species indicated that DK17 is most closely related to the polychlorinated-biphenyl-degrading *Rhodococcus jostii* RHA1 among 14 *Rhodococcus* species whose genome sequences are available in the NCBI database. Sequence alignment between ORFs from DK17 and RHA1 using blastp (protein-protein BLAST) revealed that 75% of the ORFs (6,927) share over 90% identity with each other while 9% of the ORFs (824) are present only in DK17. Among the remaining 1,515 ORFs, 245, 130, 130, 99, and 491 ORFs share identity ranging from 80 to 89, 70 to 79, 60 to 69, 50 to 59, and less than 50% with those of RHA1, respectively, while the remaining 420 have nonreciprocal matches. The DK17-specific ORFs are not randomly distributed over the genome, but rather, they occur in groups or clusters. Analysis of the DK17-specific ORFs indicates that many encode proteins likely to be involved in catabolism or transport of organics, including aromatic hydrocarbons.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. AJLQ00000000. The version described in this paper is the first version, AJLQ01000000.

ACKNOWLEDGMENTS

This work was supported by the 21C Frontier Microbial Genomics and Applications Center Program and by Basic Science Research Program (grant 2011-0001111) through the National Research Foundation of Korea. Both are funded by the Ministry of Education, Science, and Technology of the Republic of Korea. D.K. acknowledges the support of KOPRI under projects PE12040 and PM11090. G.J.Z. acknowledges the support of NIEHS grant P42-ES004911 under the Superfund Research Program. M.Y. is a recipient of the Brain Korea 21 scholarship.

REFERENCES

1. **Aziz RK, et al.** 2008. The RAST server: rapid annotations using subsystems technology. BMC Genomics **9**:75. doi:10.1186/1471-2164-9-75.
2. **Choi KY, et al.** 2005. Molecular and biochemical analysis of phthalate and terephthalate degradation by *Rhodococcus* sp. strain DK17. FEMS Microbiol. Lett. **252**:207–213.
3. **Kim D, et al.** 2002. Monocyclic aromatic hydrocarbon degradation by *Rhodococcus* sp. strain DK17. Appl. Environ. Microbiol. **68**:3270–3278.
4. **Kim D, et al.** 2004. Identification of a novel dioxygenase involved in metabolism of *o*-xylene, toluene, and ethylbenzene by *Rhodococcus* sp. strain DK17. Appl. Environ. Microbiol. **70**:7086–7092.
5. **Kim D, et al.** 2010. Aromatic hydroxylation of indan by *o*-xylene-degrading *Rhodococcus* sp. strain DK17. Appl. Environ. Microbiol. **76**:375–377.
6. **Kim D, et al.** 2011. Differential degradation of bicyclics with aromatic and alicyclic rings by *Rhodococcus* sp. strain DK17. Appl. Environ. Microbiol. **77**:8280–8287.
7. **Martinkova L, Uhnakova B, Patek M, Nesvera J, Kren V.** 2009. Biodegradation potential of the genus *Rhodococcus*. Environ. Int. **35**:162–177.
8. **Tatusov RL, et al.** 2003. The COG database: an updated version includes eukaryotes. BMC Bioinformatics **4**:41. doi:10.1186/1471-2105-4-41.

Received 14 May 2012 **Accepted** 31 May 2012

Address correspondence to Eungbin Kim, eungbin@yonsei.ac.kr.

* Present address: Ki Young Choi, Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA.

doi:10.1128/JB.00844-12