Genome sequence of *Rhodococcus* sp. R04, a polychlorinated biphenyl biodegrader

Xiuqing Yang¹*, Rui Xue¹, Chong Shen¹, Shuren Li¹, Chong Gao¹, Qi Wang², Xiaoxia Zhao³

¹Shanxi University, Institute of Biotechnology, Key Laboratory of Chemical Biology and Molecular Engineering of Ministry of Education, 030006 Taiyuan, PR China

²Shanxi University, School of Life Science, 030006 Taiyuan, PR China

³Shanxi University, Institute of Applied Chemistry, 030006 Taiyuan, PR China

*Corresponding author. Tel.: +86 351 7017661; fax: +86 351 7010215.
E-mail address: xiuqyang@sxu.edu.cn (X.Q. Yang).
Abstract:

The genus Rhodococcus has proved to be a promising option for the clean-up of polluted sites and application of a microbial biocatalyst. *Rhodococcus* sp. strain R04, isolated from oil contaminated soil, could biodegrade polychlorinated biphenyls. Here we report the draft genome sequence of *Rhodococcus* sp. strain R04, which could be used to predict genes for xenobiotic biodegradation and provide important insights into the applications of this strain.
The genus Rhodococcus is a very diverse group of bacteria that possesses the ability to degrade a large number of organic compounds, including some of the most difficult compounds with respect to recalcitrance and toxicity. Several strains belonging to the genus Rhodococcus have been isolated from various contaminated environments (1, 3, 10, 11, 12), which are proved to be ideal candidates for enhancing the bioremediation of contaminated sites and a wide range of biotransformations, such as steroid modifications, enantioselective synthesis and the production of amides from nitriles (4, 14).

*Rhodococcus* sp. R04 was isolated from oil contaminated soil in northern China. Strain R04 was able to biodegrade polychlorinated biphenyls not only via ring cleavage but also through dechlorination (15). In addition, it metabolized phenol, benzoate and 2-nitropropane when provided as the sole source of carbon and energy, and cometabolized pentachlorophenol, dibenzofuran, benzothiophene and atrazine. It is suggested that the strain R04 will be potentially useful in the biotreatment of wastewater and bioremediation of contaminated soils.

The genome sequence of R04 was determined by using the high-throughput Solexa sequencing technology (Illumina GA2x) in Shenzhen, China. WGS sequence data for 825 Mb, giving approximately 92-fold genome coverage, were generated and assembled into 2548 contigs using SOAPdenovo v.1.04 (7). Furthermore, the contigs were
joined into 110 scaffolds (>1 kb in size) using paired-end information. The genome sequence analysis of R04 showed a genome size of 9,125,386 bp with a mean GC content of 69.62%. Annotation of the open reading frames was performed using Glimmer v.3.0 (2) and by comparison with the corresponding data from the COG, KEGG, SwissProt, TrEMBL, and NR databases. tRNA and rRNA genes were identified by tRNAscan and RNAmmer, respectively (6, 8). There were 9,318 CDSs with an average length of 826 bp, 99 tRNAs and 20 rRNAs. At least 82 genes were found to be potentially involved in xenobiotic metabolism. Among these genes, four extradiol dioxygenase genes (bphC) and two hydrolase genes (bphD) were involved in PCB degradation by R04. In addition, R04 genome contains at least six ring-hydroxylating dioxygenase genes and fifteen cytochromes P450 genes, which enables the strain R04 to adapt to catabolize a large number of substrates. Differed from Rhodococcus jostii RHA1 (formerly named by Rhodococcus sp. RHA1), whose genome holds two set of polychlorinated biphenyl transformation systems (5), Rhodococcus sp. R04 only contains a bph gene cluster consisting of bphA1A2A3A4, bphB, bphC and bphD (16). In the strain R04, ketosteroid-9-α-hydrolase, 3-ketosteroid-δ-dehydrogenase, 3-ketosteroid-1-dehydrogenase and steroid δ-isomerase are involved in the transformation and metabolism of steroid compounds, they share 69-96% of identity with those of
Rhodococcus jostii RHA1, Rhodococcus rhodochrous and Rhodococcus erythropolis (9, 10, 13). It is suggested that Rhodococcus sp. R04 is a potential candidate for the industrial production of bioactive steroid compounds.

Nucleotide sequence accession numbers

This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession AFAQ00000000. The version described in this paper is the first version, AFAQ01000000.

Acknowledgements

This research was supported by the National Natural Science Foundation of China (30800030), the Young Science Foundation of Shanxi Province (207021030) and the Natural Science Foundation of Shanxi Province (2007031003).

References:


10. Morii, S., et al. 1998. 3-Ketosteroid-delta1-dehydrogenase of *Rhodococcus rhodochrous*: sequencing of the genomic DNA and...


