Genome Sequence of the Thermostable-Agarase-Producing Marine Bacterium *Catenovulum agarivorans* YM01^T, Which Reveals the Presence of a Series of Agarase-Encoding Genes

Xiaochong Shi, Min Yu, Shulin Yan, Sujie Dong, and Xiao-Hua Zhang

College of Marine Life Sciences, Ocean University of China, Qingdao, China

Marine bacterium *Catenovulum agarivorans* YM01^T can produce highly thermostable agarases. The draft genome of YM01^T is about 5.36 Mb and harbors approximately 4,913 genes, including 15 agarase (2 α-agarase and 13 β-agarase)-encoding genes, which will provide references to functional characterization of various agarases from marine bacteria.

The Gram-negative, rod-shaped, strictly aerobic, peritrichously flagellated, chain-forming, agar-hydrolyzing marine bacterium *Catenovulum agarivorans* YM01^T was recently isolated from seawater of the Yellow Sea in the coastal region of Qingdao, China, and was identified as a novel genus and species (5). Analysis of the 16S rRNA gene sequence revealed that the closest neighbors of YM01^T were the genera *Glaciecola* (88.4 to 91.0%), *Colwellia* (88.9 to 90.6%), *Vibrio* (88.9 to 90.4%), *Aestuariibacter* (89.3 to 90.4%), *Bowmanella* (90.1 to 90.3%), and *Alteromonas* (88.7 to 90.1%) in the class of *Gammaproteobacteria*. Further studies showed that YM01^T could produce thermostable agarases. After treatment at 100°C for 2 h, the crude extraction of agarases resulted in an agar-hydrolyzing capacity of 33.3% (data not shown). Thermostable agarase has various advantages over mesothermal agarase in industry applications. Here we report a draft genome sequence for *C. agarivorans* YM01^T for identification of thermostable-agarase-producing genes.

The Illumina HiSeq2000 sequencing approach with a 500-bp paired-end library was used for determining the *C. agarivorans* YM01^T genome, which achieved about 61.6-fold coverage. The reads were assembled using SOAPdenovo assembler software, and putative coding sequences (CDSs) were identified with Glimmer 3.0 (1). tRNA genes were directly predicted with tRNAscan-SE (2). Annotation was performed using the KEGG, COG, Swiss-Prot, TrEMBL, and NCBI nonredundant (NR) databases.

The draft genome of *C. agarivorans* YM01^T (about 4.87 Mb) contains 160 contigs, ranging from 59 bp to 216,200 bp (the N50 and N90 contig sizes are 82,176 bp and 20,773 bp, respectively), which can be assembled into 73 scaffolds ranging from 532 bp to 424,499 bp (N50 and N90 scaffold sizes are 212,606 bp and 59,093 bp, respectively). The G+C content was 40.04%. A total of 4,215 protein-encoding genes (89.85% among the whole genome) and 57 tRNA-containing genes were predicted in the draft genome. Among the predicted genes, 2,342 (55.6%) are assigned to COG categories (4) and 2,566 (60.9%), 2,074 (49.2%), 1,053 (25.0%), and 991 (23.5%) genes are found within the Swiss-Prot, KEGG, TrEMBL and NR databases, respectively.

More importantly, we identified 15 complete coding sequences for agarases (including 2 α-agarases and 13 β-agarases) from the *C. agarivorans* YM01^T genome, and they could be used for pro-

duction of recombinant agar-degrading agarases for various commercial applications. Until now, only one complete genome sequence of an agar-degrading marine bacterium, *Alteromonas* sp. strain S89, which showed 89.0% 16S rRNA gene similarity to *C. agarivorans* YM01^T, has been reported (3), and the results of genome annotation identified 6 agarase-encoding genes. Finally, genome sequencing of *C. agarivorans* YM01^T provides references for functional characterization of various agarases from marine bacteria.

Nucleotide sequence accession numbers. This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. AJWM00000000. The version described in this paper is the first version, AJWM01000000.

REFERENCES

1. **Delcher AL, Harmon D, Kasif S, White O, Salzberg SL.** 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Res. **27**:4636–4641.
2. **Lowe TM, Eddy SR.** 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. **25**:955–964.
3. **Oh C, et al.** 2011. Complete genome sequence of the agarase-producing marine bacterium strain S89, representing a novel species of the genus *Alteromonas*. J. Bacteriol. **193**:5538.
4. **Tatusov RL, Galperin MY, Natale DA, Koonin EV.** 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. **28**:33–36.
5. **Yan SL, Yu M, Wang Y, Shen C, Zhang X-H.** 2011. *Catenovulum agarivorans* gen. nov. sp. nov., a peritrichously flagellated, catenated, agar-hydrolyzing γ-proteobacterium isolated from the Yellow Sea. Int. J. Syst. Evol. Microbiol. **61**:2866–2873.