JOURNAL OF BACTERIOLOGY, Sept. 2011, p. 4557–4558 Vol. 193, No. 17
0021-9193/11/$12.00 doi:10.1128/JB.05513-11

Draft Genome Sequence of the Novel Agar-Digesting Marine Bacterium HQM9▽

Zongjun Du,[1,3]† Zhewen Zhang,[2]† Tingting Miao,[1] Jiayan Wu,[2] Guoqiang Lü,[1] Jun Yu,[2] Jingfa Xiao,[2]* and Guanjun Chen[1,3]*

College of Marine Science, Shandong University at Weihai, Weihai 264209, China[1]; CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100029, China[2]; and State Key Laboratory of Microbial Technology, College of Life Science, Shandong University, Jinan 250100, China[3]

Received 15 June 2011/Accepted 22 June 2011

Strain HQM9, an aerobic, rod-shaped marine bacterium from red algae, can produce agarases and liquefy solid plating media efficiently when agar is used as a coagulant. Here we report the draft genome sequence and the initial findings from a preliminary analysis of strain HQM9, which should be a novel species of *Flavobacteriaceae*.

Strain HQM9, which was isolated from the surfaces of red algae, is a yellow-pigmented, aerobic, Gram-negative, agar-degrading bacterium that represents a novel species in the family *Flavobacteriaceae*, based on its 16S rRNA gene sequence. The marine flavobacteria are known for producing enzymes that degrade polysaccharides such as agar, laminarin, xylan, fucoidan, and carrageenan from micro- or macroalgae (3, 5, 11, 13, 21). The agar-degrading bacteria play an important role in the marine carbon cycle involving the breakdown of agar and other sulfated galactans, which form a significant component of the cell walls of red and green algae, the egg jelly coating of certain sea urchin species, and the outer tunics of ascidians (2, 19). Over the past couple of years, the genome sequences of *Flavobacteriaceae* family members *Flavobacterium psychrophilum* JIP02/86 (8), *Robiginitalea biformata* HTCC2501 (17), *Capnocytophaga ochracea* DSM 7271 (15), *Zunongwangia profunda* SM-A87 (20), and *Gramella forsetii* KT0803 (4) have been published.

The genome of HQM9 was sequenced with a combined strategy of 454 genome sequencer FLX (454 GS FLX) sequencing and Illumina paired-end sequencing at the Beijing Institute of Genomics. The 454 GS FLX sequencing achieved about 21-fold coverage, and 498-fold coverage of reads was achieved by Illumina paired-end sequencing. The draft genome (about 4 Mbp) contains 183 contigs, which can be assembled into 74 scaffolds. Scaffold N50 is 440,279 bp. The GC content of the HQM9 draft genome is 33.2%. We predicted the tRNA genes by tRNAscan-SE (14). Ribosomal RNAs were found by BLAST searching against the Rfam database (10). Open reading frames (ORFs) were identified by using Glimmer 3.0 (7) and GeneMarkS (6). All predicted ORFs were then annotated by BLAST (1), InterPro (16), and KEGG.

The draft genome contains 3,971 protein-coding genes, 2 rRNA operons, and 37 tRNA genes. Three thousand five hundred nineteen predicted protein-coding genes have homologs in GenBank databases of nonredundant protein sequences (E value $< 1e-5$). Approximately 7% of HQM9 genes have similarity (identity of $\geq 30\%$) to those of *Croceibacter atlanticus* HTCC2559 (18), which also belongs to the *Flavobacteriaceae* family and for which the complete genome has been sequenced. Notably, 34 agarase genes, the most agarase genes detected in one bacterial genome so far, were found in the HQM9 draft genome. The agarases can be grouped into α-agarases and β-agarases according to the cleavage pattern presently known (9). The 34 agarases of HQM9 all belong to the β-agarase group, based on sequence similarity. Furthermore, the catalytic domains of β-agarases have been classified into three glycoside hydrolase (GH) families, i.e., GH-16, GH-50, and GH-86 (9). Of these 34 agarases, 14 belong to the GH-16 family, 6 belong to GH-86, and only 2 belong to GH-50 (http://www.cazy.org/glycoside-hydrolases.html). Agarase can catalyze the degradation of agarose polysaccharide into neoagaro-oligosaccharides by cleavage of the β-1,4 linkages (12, 22) and can help the bacterium get enough nutrients from the algae. Many other enzymes for degradation were also identified. Fifty-seven peptidases for digesting proteins and 14 glycoside hydrolases for digesting polysaccharides were predicted. According to the KEGG pathway analysis, most genes encode the HQM9 proteins for glycolysis, the citrate cycle (TCA cycle), the pentose phosphate pathway, galactose metabolism, and fatty acid metabolism. These metabolism pathways may provide enough energy to HQM9 for adapting to the complicated and changeable marine environment.

A more specific analysis of strain HQM9 will be reported in a future publication.

Nucleotide sequence accession number. The draft genome sequence of HQM9 is available in GenBank under accession number AFPB00000000.

* Corresponding author. Mailing address for Guanjun Chen: State Key Laboratory of Microbial Technology, College of Life Science, Shandong University, Jinan 250100, China. Phone: 86-531-88366202. Fax: 86-531-88364430. E-mail: guanjun@sdu.edu.cn. Mailing address for Jingfa Xiao: CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100029, China. Phone: 86-10-82995384. Fax: 86-10-82995401. E-mail: xiaojingfa@big.ac.cn.

† The first two authors contributed equally to this work.

▽ Published ahead of print on 1 July 2011.

This work was supported by the National Science Foundation of China (40730847), the Independent Innovation Foundation of Shandong University (IIFSDU), and the Special Fund for Shangdong Provincial Post-Doctoral Innovative Programs (201003058).

REFERENCES

1. **Altschul, S. F., et al.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
2. **Armisen, R., and F. Galactas.** 1987. Production, properties and uses of agar, p. 1–57. *In* D. J. McHugh (ed.), Production and utilization of products from commercial seaweeds. FAO Fisheries Technical Paper 288. Fisheries and Aquaculture Department, Food and Agriculture Organization of the United Nations, New York, NY.
3. **Barbeyron, T., S. L'Haridon, E. Corre, B. Kloareg, and P. Potin.** 2001. *Zobellia galactanovorans* gen. nov., sp. nov., a marine species of *Flavobacteriaceae* isolated from a red alga, and classification of [*Cytophaga*] *uliginosa* (ZoBell and Upham 1944) Reichenbach 1989 as *Zobellia uliginosa* gen. nov., comb. nov. Int. J. Syst. Evol. Microbiol. **51:**985–997.
4. **Bauer, M., et al.** 2006. Whole genome analysis of the marine *Bacteroidetes* '*Gramella forsetii*' reveals adaptations to degradation of polymeric organic matter. Environ. Microbiol. **8:**2201–2213.
5. **Bernardet, J.-F., and Y. Nakagawa.** 2006. An introduction to the family Flavobacteriaceae. Prokaryotes **7:**455–480.
6. **Besemer, J., A. Lomsadze, and M. Borodovsky.** 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. Nucleic Acids Res. **29:**2607–2618.
7. **Delcher, A. L., K. A. Bratke, E. C. Powers, and S. L. Salzberg.** 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics **23:**673–679.
8. **Duchaud, E., et al.** 2007. Complete genome sequence of the fish pathogen Flavobacterium psychrophilum. Nat. Biotechnol. **25:**763–769.
9. **Fu, X. T., and S. M. Kim.** 2010. Agarase: review of major sources, categories, purification method, enzyme characteristics and applications. Mar. Drugs **8:**200–218.
10. **Griffiths-Jones, S., A. Bateman, M. Marshall, A. Khanna, and S. R. Eddy.** 2003. Rfam: an RNA family database. Nucleic Acids Res. **31:**439–441.
11. **Humphry, D. R., A. George, G. W. Black, and S. P. Cummings.** 2001. *Flavobacterium frigidarium* sp. nov., an aerobic, psychrophilic, xylanolytic and laminarinolytic bacterium from Antarctica. Int. J. Syst. Evol. Microbiol. **51:**1235–1243.
12. **Jam, M., et al.** 2005. The endo-β-agarases AgaA and AgaB from the marine bacterium *Zobellia galactanivorans*: two paralogue enzymes with different molecular organizations and catalytic behaviours. Biochem. J. **385:**703–713.
13. **Johansen, J. E., P. Nielsen, and C. Sjoholm.** 1999. Description of *Cellulophaga baltica* gen. nov., sp. nov. and *Cellulophaga fucicola* gen. nov., sp. nov. and reclassification of [*Cytophaga*] *lytica* to *Cellulophaga lytica* gen. nov., comb. nov. Int. J. Syst. Bacteriol. **49:**1231–1240.
14. **Lowe, T. M., and S. R. Eddy.** 1997. tRNAscan-SE: a program for improved detection of tRNA genes in genomic sequence. Nucleic Acids Res. **25:**955–964.
15. **Mavrommatis, K., et al.** 2009. Complete genome sequence of *Capnocytophaga ochracea* type strain (VPI 2845). Stand. Genomic Sci. **1:**101–109.
16. **Mulder, N., and R. Apweiler.** 2007. InterPro and InterProScan: tools for protein sequence classification and comparison. Methods Mol. Biol. **396:** 59–70.
17. **Oh, H. M., et al.** 2009. Complete genome sequence of *Robiginitalea biformata* HTCC2501. J. Bacteriol. **191:**7144–7145.
18. **Oh, H. M., I. Kang, S. Ferriera, S. J. Giovannoni, and J. C. Cho.** 2010. Complete genome sequence of *Croceibacter atlanticus* HTCC2559T. J. Bacteriol. **192:**4796–4797.
19. **Pomin, V. H.** 2010. Structural and functional insights into sulfated galactans: a systematic review. Glycoconj. J. **27:**1–12.
20. **Qin, Q. L., et al.** 2010. The complete genome of *Zunongwangia profunda* SM-A87 reveals its adaptation to the deep-sea environment and ecological role in sedimentary organic nitrogen degradation. BMC Genomics **11:**247.
21. **Sakai, T., H. Kimura, and I. Kato.** 2002. A marine strain of flavobacteriaceae utilizes brown seaweed fucoidan. Mar. Biotechnol. (NY) **4:**399–405.
22. **Zhang, W. W., and L. Sun.** 2007. Cloning, characterization, and molecular application of a beta-agarase gene from *Vibrio* sp. strain V134. Appl. Environ. Microbiol. **73:**2825–2831.