

Genome Sequence of Strain IMCC3088, a Proteorhodopsin-Containing Marine Bacterium Belonging to the OM60/NOR5 Clade[▽]

Yoonra Jang, Hyun-Myung Oh, Innam Kang, Kiyoun Lee, Seung-Jo Yang, and Jang-Cheon Cho*

Division of Biology and Ocean Sciences, Inha University, Incheon 402-751, Republic of Korea

Received 18 April 2011/Accepted 25 April 2011

Strain IMCC3088, cultivated from the Yellow Sea, is a novel isolate belonging to the OM60/NOR5 clade and is closely related to clone OM241, *Congregibacter litoralis*, and strain HTCC2080. Here, the genome sequence of strain IMCC3088 is presented, showing the absence of photosynthetic gene clusters and the presence of proteorhodopsin.

The OM60/NOR5 clade, part of the oligotrophic marine *Gammaproteobacteria* group, comprises up to 11% of the bacterial communities on the ocean surface (3, 5, 6, 9). This important marine clade includes *Congregibacter litoralis* (6, 7) and extinction cultures designated HTCC2080, HTCC2148, and HTCC2246 (3, 4), in addition to several uncharacterized marine oligotrophic bacteria. Full genomic analyses of two representative strains (HTCC2080 and *C. litoralis* KT71) in the OM60/NOR5 clade have shown that they are aerobic anoxygenic phototrophs (AAnP) harboring photosynthetic gene clusters (*puf*) (4, 6). However, genome analyses of strain HTCC2148 (8), a member of the NOR5-8 subclade (9), showed the absence of photosynthetic gene clusters in the genome, indicating that there would be more diverse metabolism other than AAnP in the OM60/NOR5 clade.

Strain IMCC3088 was isolated from a surface seawater sample collected from the Yellow Sea off the coast of Incheon, South Korea, by using standard dilution techniques on an oligotrophic medium, marine 1/10R2A. Phylogenetic analysis based on 16S rRNA gene sequences showed that strain IMCC3088 belongs to the NOR5-4 subclade (9) of the OM60/NOR5 clade, which contains OM241 (U70702) and DS029 (DQ234113). The strain is also closely related to cultured marine isolates of *C. litoralis* KT71 (94.8%), *Haliea rubra* CM41-15a (94.7%), and strain HTCC2080 (93.9%). Strain IMCC3088 was chosen for genome sequencing to reveal diverse metabolic potential in the OM60/NOR5 clade.

Whole-genome shotgun sequencing of strain IMCC3088 was performed by Macrogen, Inc. (Seoul, South Korea), with library construction using 5 µg genomic DNA based on Solexa paired-end sequencing technology (2). A total of 7,015,036 short reads were generated, reaching a depth of 86-fold genome coverage, and assembled into 183 contigs (AEIG01000001 to AEIG01000183) by using Velvet (10). The unclosed draft genome of IMCC3088 has 3,045,708 bases with a mean G+C content of 51.64%. Annotation of

the open reading frames (ORFs) and noncoding RNAs was performed by using the RAST (rapid annotation using subsystem technology) server (1), and predicted ORFs were cross-checked with the COG, KEGG, and SwissProt databases. The IMCC3088 genome has 2,858 predicted ORFs, one copy each of the 5S, 16S, and 23S rRNA genes, and 34 predicted tRNAs.

In contrast to the existence of a *puf* operon in *C. litoralis* KT71 and HTCC2080, there are no predicted *puf* genes in the genome of IMCC3088. Instead, the IMCC3088 genome is predicted to encode proteorhodopsin, followed by the gene clusters for the biosynthesis of the retinal chromophore of rhodopsin (*crtEIBY*), suggesting that proteorhodopsin is functional. Strain IMCC3088 could be the first bacterium that harbors proteorhodopsin among the members of the OM60/NOR5 clade, indicating the existence of diverse metabolic potential in the members of this marine clade. The predicted ORFs of the IMCC3088 genome also include genes encoding enzymes for the Entner-Doudoroff pathway, the pentose phosphate pathway, ABC transporters for inorganic ions and amino acids, sulfur oxidation (*soxCDXYZAB*), aromatic compound degradation, and C₁ compound metabolism.

Nucleotide sequence accession number. The genome sequence of strain IMCC3088 is available in GenBank under accession number AEIG00000000.

This study was supported by the 21C Frontier Program of Microbial Genomics and Applications and by the Korea Research Foundation Grant funded by the Korean Government (MEST, KRF-2008-521-C00263).

REFERENCES

1. Aziz, R., et al. 2008. The RAST server: rapid annotations using subsystems technology. *BMC Genomics* **9**:75.
2. Bentley, D. R., et al. 2008. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456**:53–59.
3. Cho, J.-C., and S. J. Giovannoni. 2004. Cultivation and growth characteristics of a diverse group of oligotrophic marine *Gammaproteobacteria*. *Appl. Environ. Microbiol.* **70**:432–440.
4. Cho, J.-C., et al. 2007. Polyphyletic photosynthetic reaction center genes in oligotrophic marine *Gammaproteobacteria*. *Environ. Microbiol.* **9**:1456–1463.
5. Eilers, H., J. Pernthaler, F. O. Glöckner, and R. Amann. 2000. Culturability and in situ abundance of pelagic bacteria from the North Sea. *Appl. Environ. Microbiol.* **66**:3044–3051.
6. Fuchs, B. M., et al. 2007. Characterization of a marine gammaproteobacte-

* Corresponding author. Mailing address: Division of Biology and Ocean Sciences, Inha University, Incheon 402-751, Republic of Korea. Phone: 82-32-860-7711. Fax: 82-32-232-0541. E-mail: chojc@inha.ac.kr.

[▽] Published ahead of print on 6 May 2011.

- rium capable of aerobic anoxygenic photosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **104**:2891–2896.
7. **Spring, S., H. Lunsdorf, B. M. Fuchs, and B. J. Tindall.** 2009. The photosynthetic apparatus and its regulation in the aerobic gammaproteobacterium *Congregibacter litoralis* gen. nov., sp. nov. *PLoS One* **4**:e4866.
8. **Thrash, J. C., et al.** 2010. Genome sequences of strains HTCC2148 and HTCC2080, belonging to the OM60/NOR5 clade of the *Gammaproteobacteria*. *J. Bacteriol.* **192**:3842–3843.
9. **Yan, S., et al.** 2009. Biogeography and phylogeny of the NOR5/OM60 clade of Gammaproteobacteria. *Syst. Appl. Microbiol.* **32**:124–139.
10. **Zerbino, D. R., and E. Birney.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* **18**:821–829.