

Genome Sequences of Strains HTCC2148 and HTCC2080, Belonging to the OM60/NOR5 Clade of the *Gammaproteobacteria*[▽]

J. Cameron Thrash,^{1†} Jang-Cheon Cho,^{2†} Steve Ferriera,³ Justin Johnson,³
Kevin L. Vergin,¹ and Stephen J. Giovannoni^{1*}

Department of Microbiology, Oregon State University, Corvallis, Oregon 97331¹; Division of Biology and Ocean Sciences, Inha University, Incheon 402-751, Republic of Korea²; and J. Craig Venter Institute, 9704 Medical Center Drive, Rockville, Maryland 20850³

Received 4 May 2010/Accepted 5 May 2010

Organisms in the OM60/NOR5 clade of the *Gammaproteobacteria* are ubiquitous in the world's oceans and can make up as much as 11% of bacterial cells in certain areas. Isolated from coastal Oregon water, *Gammaproteobacteria* HTCC2148 and HTCC2080 are two members of this important clade. Here we present the genome sequences of the OM60 *Gammaproteobacteria* HTCC2148 and HTCC2080.

Strains HTCC2148 and HTCC2080 are members of the important OM60/NOR5 clade of the oligotrophic marine *Gammaproteobacteria* (OMG), which includes the aerobic anoxygenic phototroph *Congregibacter litoralis* KT71 in addition to many species of chemoheterotrophs (6, 7). Organisms from this clade are detected globally throughout the euphotic zone (1, 5, 14) and can make up a significant portion ($\leq 11\%$) of bacterial cells in some locations (1, 6, 12, 14). Consistent with this distribution, HTCC2148 and HTCC2080 were isolated from Oregon coastal water collected at hydrostation NH5 from a depth of 10 m (2). Using dilution-to-extinction culturing methods previously described (3, 13), these organisms were cultivated in low-nutrient heterotrophic medium (LNHM) (2) and selected for sequencing due to their phylogenetic affiliation.

As part of the Moore Foundation Microbial Genome Sequencing Project (<http://www.moore.org/microgenome>), the genomes were shotgun sequenced by the J. Craig Venter Institute, and the draft genomes of HTCC2148 and HTCC2080 containing 75 (ABXQ01000001 to ABXQ01000075) and 25 (AAVV01000001 to AAVV01000025) contigs, respectively, were annotated with the GenDB annotation application program (9) at the Center for Genome Research and Biocomputing at Oregon State University, similarly to that previously described (10, 11). Autoannotation was completed by merging results from the Glimmer 2.0 modeling package (4) and contrasting to Basic Local Alignment Search Tool for Proteins (BLASTp). The protein set was additionally searched against the KEGG, SwissProt, Clusters of Orthologous Groups (COG), Pfam, and Interpro protein databases to annotate EC number, gene call, and gene description. These GenDB annotations predicted major metabolic pathways and biosynthesis of amino acids, vitamins, and growth factors, with manual BLAST searches run to confirm autoannotations when neces-

sary. G+C mol% measurements were computed using the genome sequence and calculated using Practical Extraction and Reporting Language (PERL). Complementary analysis was performed through the Joint Genome Institute IMG/M website (<http://img.jgi.doe.gov/cgi-bin/pub/main.cgi>) (8).

The draft genomes of HTCC2148 and HTCC2080 comprised 4,310,225 and 3,576,081 bases, 3,866 and 3,229 predicted open reading frames (ORFs), and 52.96% and 51.82% G+C contents, respectively. The HTCC2148 genome is predicted to have 39 tRNA genes and two each of the 5S, 16S, and 23S rRNA genes; that of HTCC2080 is predicted to have 41 tRNA genes and one each of the 5S, 16S, and 23S rRNA genes. Both organisms are predicted to have complete glycolysis and Entner-Doudoroff pathways, a complete tricarboxylic acid (TCA) cycle, and the ability to make all 20 essential amino acids and some vitamins and cofactors. These organisms may be deficient in B6 and thiamine production, however, and HTCC2080 has no predicted genes for biotin production. Both genomes have putative catalase and alcohol and formate dehydrogenase genes. HTCC2080 contains some predicted phototrophy genes, including a predicted reaction-center H chain, and both contain predicted phosphoenolpyruvate carboxylase genes, indicating the potential for autotrophy in these organisms.

Nucleotide sequence accession numbers. The draft genome sequences of *Gammaproteobacteria* HTCC2148 and HTCC2080 have the accession numbers ABXQ00000000 and AAVV00000000, respectively, in GenBank. Data annotated independently with GenDB are also accessible in the Marine Microbial Genomics database at Oregon State University (<http://bioinfo.cgrb.oregonstate.edu/microbes/>).

Sequencing, assembly, and annotation and data analysis were supported by the Gordon and Betty Moore Foundation Marine Microbiology Initiative as part of its Marine Microbial Sequencing Project (<http://www.moore.org/marine-micro.aspx>) and by an individual investigator award to S.J.G.

We also thank Scott Givan at the Oregon State University Center for Genome Research and Biocomputing for his work in developing the genome annotation software.

REFERENCES

1. Cho, J.-C., M. D. Stapels, R. M. Morris, K. L. Vergin, M. S. Schwalbach, S. A. Givan, D. F. Barofsky, and S. J. Giovannoni. 2007. Polyphyletic pho-

* Corresponding author. Mailing address: Department of Microbiology, Oregon State University, Corvallis, OR 97331. Phone: (541) 737-1835. Fax: (541) 737-0496. E-mail: steve.giovannoni@oregonstate.edu.

† These authors contributed equally to this work.

[▽] Published ahead of print on 14 May 2010.

- tosynthetic reaction centre genes in oligotrophic marine *Gammaproteobacteria*. *Environ. Microbiol.* **9**:1456–1463.
2. **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.
 3. **Connon, S. A., and S. J. Giovannoni.** 2002. High-throughput methods for culturing microorganisms in very-low-nutrient media yield diverse new marine isolates. *Appl. Environ. Microbiol.* **68**:3878–3885.
 4. **Delcher, A. L., D. Harmon, S. Kasif, O. White, and S. L. Salzberg.** 1999. Improved microbial gene identification with GLIMMER. *Nucleic Acids Res.* **27**:4636–4641.
 5. **DeLong, E., C. M. Preston, T. Mincer, V. Rich, S. J. Hallam, N. Frigaard, A. Martinez, M. B. Sullivan, R. Edwards, B. R. Brito, S. W. Chisholm, and D. M. Karl.** 2006. Community genomics among stratified microbial assemblages in the ocean's interior. *Science* **311**:496–503.
 6. **Eilers, H., J. Pernthaler, J. Peplies, F. Glockner, G. Gerdt, and R. Amann.** 2001. Isolation of novel pelagic bacteria from the German bight and their seasonal contributions to surface picoplankton. *Appl. Environ. Microbiol.* **67**:5134–5142.
 7. **Fuchs, B. M., S. Spring, H. Teeling, C. Quast, J. Wulf, M. Schattenhof, S. Yan, S. Ferriera, J. Johnson, F. O. Glockner, and R. Amann.** 2007. Characterization of a marine gammaproteobacterium capable of aerobic anoxygenic photosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **104**:2891–2896.
 8. **Markowitz, V. M., N. N. Ivanova, E. Szeto, K. Palaniappan, K. Chu, D. Dalevi, I.-M. A. Chen, Y. Grechkin, I. Dubchak, I. Anderson, A. Lykidis, K. Mavromatis, P. Hugenholtz, and N. C. Kyrpides.** 2008. IMG/M: a data management and analysis system for metagenomes. *Nucleic Acids Res.* **36**:D534–D538.
 9. **Meyer, F., A. Goesmann, A. McHardy, D. Bartels, T. Bekel, J. Clausen, J. Kalinowski, B. Linke, O. Rupp, R. Giegerich, and A. Puhler.** 2003. GenDB—an open source genome annotation system for prokaryote genomes. *Nucleic Acids Res.* **31**:2187–2195.
 10. **Oh, H.-M., S. J. Giovannoni, S. Ferriera, J. Johnson, and J.-C. Cho.** 2009. Complete genome sequence of *Erythrobacter litoralis* HTCC2594. *J. Bacteriol.* **191**:2419–2420.
 11. **Oh, H.-M., S. J. Giovannoni, K. Lee, S. Ferriera, J. Johnson, and J.-C. Cho.** 2009. Complete genome sequence of *Robiginitalea biformata* HTCC2501. *J. Bacteriol.* **191**:7144–7145.
 12. **Pernthaler, A., and J. Pernthaler.** 2005. Diurnal variation of cell proliferation in three bacterial taxa from coastal North Sea waters. *Appl. Environ. Microbiol.* **71**:4638–4644.
 13. **Stingl, U., H. J. Tripp, and S. J. Giovannoni.** 2007. Improvements of high-throughput culturing yielded novel SAR11 strains and other abundant marine bacteria from the Oregon coast and the Bermuda Atlantic Time Series study site. *ISME J.* **1**:361–371.
 14. **Yan, S., B. M. Fuchs, S. Lenk, J. Harder, and J. Wulf.** 2009. Biogeography and phylogeny of the NOR5/OM60 clade of *Gammaproteobacteria*. *Syst. Appl. Microbiol.* **32**:124–139.