Genome Sequence of *Porticoccus hydrocarbonoclasticus* Strain MCTG13d, an Obligate Polycyclic Aromatic Hydrocarbon-Degrading Bacterium Associated with Marine Eukaryotic Phytoplankton

Tony Gutierrez,[a] William B. Whitman,[b] Marcel Huntemann,[c] Alex Copeland,[c] Amy Chen,[c] Nikos Kyrpides,[c] Victor Markowitz,[c] Manoj Pillay,[c] Natalia Ivanova,[c] Natalia Mikhailova,[c] Galina Ovchinnikova,[c] Evan Andersen,[c] Amrita Pati,[c] Dimitrios Stamatis,[c] T. B. K. Reddy,[c] Chew Yee Ngan,[c] Mansi Chovatia,[c] Chris Daum,[c] Nicole Shapiro,[c] Michael N. Cantor,[c] Tanja Woyke[c]

School of Life Sciences, Heriot-Watt University, Edinburgh, United Kingdom[a]; Department of Microbiology, University of Georgia, Athens, Georgia, USA[b]; DOE Joint Genome Institute, Walnut Creek, California, USA[c]

Porticoccus hydrocarbonoclasticus strain MCTG13d is a recently discovered bacterium that is associated with marine eukaryotic phytoplankton and that almost exclusively utilizes polycyclic aromatic hydrocarbons (PAHs) as the sole source of carbon and energy. Here, we present the genome sequence of this strain, which is 2,474,654 bp with 2,385 genes and has an average G+C content of 53.1%.

Address correspondence to Tony Gutierrez, tony.gutierrez@hw.ac.uk.

*P*orticoccus hydrocarbonoclasticus strain MCTG13d was isolated from a laboratory culture of the marine dinoflagellate *Lingulodinium polyedrum* (CCAP 1121/2) by enrichment with polycyclic aromatic hydrocarbons (PAHs) as the sole carbon source (1). The strain represents a novel species of obligate hydrocarbonoclastic marine bacteria (OHCB) that exhibit a narrow nutritional spectrum, preferring to utilize aliphatic and aromatic hydrocarbons and small organic acids (2). Notably, strain MCTG13d displays versatility for degrading three- and four-ring PAHs, consistent with the catabolic spectrum of members belonging to the obligate PAH-degrading genera *Cycloclasticus* (3) and *Neptunomonas* (4). Strain MCTG13d is a strictly aerobic and motile, rod-shaped bacterium that is associated with various species of marine diatoms and dinoflagellates found in different seas and oceans worldwide (1; T. Gutierrez, unpublished results).

Here, we report the genome sequence of *Porticoccus hydrocarbonoclasticus* strain MCTG13d. Genomic DNA was isolated, and the sequence was generated at the Department of Energy (DOE) Joint Genome Institute (JGI; Walnut Creek, CA, USA) using Pacific Biosciences (PacBio) technology. A PacBio SMRTbellTM library was constructed and sequenced on the PacBio RS platform, which generated 189,901 filtered subreads totaling 628.8 Mbp. All general aspects of library construction and sequencing performed at the JGI can be found at http://www.jgi.doe.gov. The raw reads were assembled using HGAP version 2.1.1 (5). The final draft assembly produced 1 scaffold containing 1 contig totaling 2.5 Mbp and input read coverage of 291.1×.

Project information is available in the Genomes OnLine Database (6). Genes were identified using Prodigal (7), followed by a round of manual curation using GenePRIMP (8) as part of the JGI's microbial annotation pipeline (9). The predicted coding sequences were translated and used to search the National Center for Biotechnology Information (NCBI) nonredundant, UniProt, TIGRFam, Pfam, KEGG, COG, and InterPro databases. The tRNAscanSE tool (10) was used to find tRNA genes, whereas rRNA genes were found by searches against models of the rRNA genes built from SILVA (11). Other noncoding RNAs such as the RNA components of the protein secretion complex and the RNase P were identified by searching the genome for the corresponding Rfam profiles using Infernal (http://infernal.janelia.org). Additional gene prediction analysis and manual functional annotation was performed within the Integrated Microbial Genomes Expert Review (IMG ER) platform (http://img.jgi.doe.gov) developed by the JGI (12).

The complete genome sequence length was 2,474,654 bp with a G+C content of 53.1%. The genome contains 2,385 genes (2,340 protein-coding genes) with functional predictions for 2,021 of them. A total of 45 RNA genes were detected. Other genes, characteristic for the genus, are given in the IMG database (12). This genome sequence is expected to provide great insights into the unusual life style of this organism.

Nucleotide sequence accession number. The draft genome sequence of *P. hydrocarbonoclasticus* strain MCTG13d obtained in this study was deposited in GenBank as part of BioProject number PRJNA224116, with individual genome sequences submitted as whole-genome shotgun projects under the accession number JQMM00000000.

ACKNOWLEDGMENTS

REFERENCES

1. **Gutierrez T, Nichols PD, Whitman WB, Aitken MD**. 2012. *Porticoccus hydrocarbonoclasticus* sp. nov., an aromatic hydrocarbon-degrading bacterium identified in laboratory cultures of marine phytoplankton. Appl Environ Microbiol **78**:628–637. http://dx.doi.org/10.1128/AEM.06398-11.

2. **Yakimov MM, Timmis KN, Golyshin PN**. 2007. Obligate oil-degrading marine bacteria. Curr Opin Biotechnol **18**:257–266. http://dx.doi.org/10.1016/j.copbio.2007.04.006.

3. **Dyksterhouse SE, Gray JP, Herwig RP, Lara JC, Staley JT**. 1995. *Cycloclasticus pugetii* gen. nov., sp. nov., an aromatic hydrocarbon-degrading bacterium from marine sediments. Int J Syst Bacteriol **45**:116–123. http://dx.doi.org/10.1099/00207713-45-1-116.

4. **Hedlund BP, Geiselbrecht AD, Bair TJ, Staley JT**. 1999. Polycyclic aromatic hydrocarbon degradation by a new marine bacterium, *Neptunomonas naphthovorans* gen. nov., sp. nov. Appl Environ Microbiol **65**:251–259.

5. **Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J**. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods **10**:563–569. http://dx.doi.org/10.1038/nmeth.2474.

6. **Reddy TB, Thomas AD, Stamatis D, Bertsch J, Isbandi M, Jansson J, Mallajosyula J, Pagani I, Lobos EA, Kyrpides NC**. 2015. The Genomes OnLine Database (GOLD) v.5: a metadata management system based on a four level (meta)genome project classification. Nucleic Acids Res **43**:D1099–D1106. http://dx.doi.org/10.1093/nar/gku950.

7. **Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ**. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics **11**:119. http://dx.doi.org/10.1186/1471-2105-11-119.

8. **Pati A, Ivanova NN, Mikhailova N, Ovchinnikova G, Hooper SD, Lykidis A, Kyrpides NC**. 2010. GenePRIMP: a gene prediction improvement pipeline for prokaryotic genomes. Nat Methods **7**:455–457. http://dx.doi.org/10.1038/nmeth.1457.

9. **Mavromatis K, Ivanova NN, Chen IM, Szeto E, Markowitz VM, Kyrpides NC**. 2009. Standard operating procedure for the annotations of microbial genomes by the Production Genomic Facility of the DOE JGI. Stand Genomic Sci **1**:63–67. http://dx.doi.org/10.4056/sigs.632.

10. **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. http://dx.doi.org/10.1093/nar/25.5.0955.

11. **Pruesse E, Quast C, Knittel K, Fuchs B, Ludwig W, Peplies J, Glöckner FO**. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res **35**:2188–2196. http://dx.doi.org/10.1093/nar/gkm864.

12. **Markowitz VM, Mavromatis K, Ivanova NN, Chen IM, Chu K, Kyrpides NC**. 2009. IMG ER: a system for microbial genome annotation expert review and curation. Bioinformatics **25**:2271–2278. http://dx.doi.org/10.1093/bioinformatics/btp393.