

UC Merced

UC Merced Previously Published Works

Title

Genome Sequence of Polycyclovorans algicola Strain TG408, an Obligate Polycyclic Aromatic Hydrocarbon-Degrading Bacterium Associated with Marine Eukaryotic Phytoplankton

Permalink

<https://escholarship.org/uc/item/96x389f4>

Journal

Microbiology Resource Announcements, 3(2)

ISSN

2169-8287

Authors

Gutierrez, Tony
Thompson, Haydn F
Angelova, Angelina
[et al.](#)

Publication Date

2015-04-30

DOI

10.1128/genomea.00207-15

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Genome Sequence of *Polycyclovorans algicola* Strain TG408, an Obligate Polycyclic Aromatic Hydrocarbon-Degrading Bacterium Associated with Marine Eukaryotic Phytoplankton

Tony Gutierrez,^a Haydn F. Thompson,^a Angelina Angelova,^a William B. Whitman,^b Marcel Huntemann,^c Alex Copeland,^c Amy Chen,^c Nikos Kyrpides,^c Victor Markowitz,^c Krishnaveni Palaniappan,^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

***Polycyclovorans algicola* strain TG408 is a recently discovered bacterium associated with marine eukaryotic phytoplankton and exhibits the ability to utilize polycyclic aromatic hydrocarbons (PAHs) almost exclusively as sole sources of carbon and energy. Here, we present the genome sequence of this strain, which is 3,653,213 bp, with 3,477 genes and an average G+C content of 63.8%.**

Received 12 February 2015 Accepted 18 February 2015 Published 26 March 2015

Citation Gutierrez T, Thompson HF, Angelova A, Whitman WB, Huntemann M, Copeland A, Chen A, Kyrpides N, Markowitz V, Palaniappan K, Ivanova N, Mikhailova N, Ovchinnikova G, Andersen E, Pati A, Stamatis D, Reddy TBK, Ngan CY, Chovatia M, Daum C, Shapiro N, Cantor MN, Woyke T. 2015. Genome sequence of *Polycyclovorans algicola* strain TG408, an obligate polycyclic aromatic hydrocarbon-degrading bacterium associated with marine eukaryotic phytoplankton. *Genome Announc* 3(2):e00207-15. doi: 10.1128/genomeA.00207-15.

Copyright © 2015 Gutierrez et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](http://creativecommons.org/licenses/by/3.0/).

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

Polycyclovorans algicola strain TG408 was isolated from a laboratory culture of the marine diatom *Skeletonema costatum* (CCAP 1077/1C) by enrichment with polycyclic aromatic hydrocarbons (PAHs) as the sole carbon source (1). The strain represents a novel genus of obligate hydrocarbonoclastic marine bacteria (OHCB) that exhibit a narrow nutritional spectrum, preferring to utilize aliphatic and aromatic hydrocarbon compounds and small organic acids (2). Notably, strain TG408 displays versatility for degrading two- and three-ring PAHs, consistent with the catabolic spectrum of members belonging to the obligate PAH-degrading genera *Cycloclasticus* (3) and *Neptunomonas* (4). Strain TG408 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) (our unpublished data).

Here, we report the genome sequence of *P. algicola* strain TG408. Genomic DNA was isolated, and the sequence was generated at the Department of Energy (DOE) Joint Genome Institute (JGI) using the Pacific Biosciences (PacBio) technology. A PacBio SMRTbell library was constructed and sequenced on the PacBio RS platform, which generated 262,160 filtered subreads totaling 814.1 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 3.7 Mbp in size and with an input read coverage of 263.8×.

Genes were identified using Prodigal (6), followed by a round of manual curation using GenePRIMP (7) for finished genomes and draft genomes in <10 scaffolds. The predicted coding se-

quences (CDSs) were translated and used to search the National Center for Biotechnology Information (NCBI) nonredundant, UniProt, TIGRFam, Pfam, KEGG, COG, and InterPro databases. The tRNAscan-SE tool (8) was used to find tRNA genes, whereas rRNA genes were found by searches against models of the rRNA genes built from SILVA (9). 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 were performed within the Integrated Microbial Genomes (IMG) platform (<http://img.jgi.doe.gov>) developed by the Joint Genome Institute, Walnut Creek, CA, USA (10).

The complete genome sequence length was 3,653,213 bp, with a G+C content of 63.8%. The genome contains 3,477 genes (3,413 protein-coding genes), with function predictions for 2,818 of them. A total of 64 RNA genes were detected. Other genes that are characteristic for the genus are given in the IMG database (10).

Nucleotide sequence accession number. The draft genome sequence of *P. algicola* strain TG408 obtained in this study was deposited in GenBank as part of BioProject no. PRJNA224116, with individual genome sequences submitted as whole-genome shotgun projects under the accession no. JOMH00000000.

ACKNOWLEDGMENTS

T.G. was supported by a Marie Curie International Outgoing Fellowship (PIOF-GA-2008-220129) within the 7th European Community Framework Program.

The work was conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, which is sup-

ported by the Office of Science of the U.S. Department of Energy under contract no. DE-AC02-05CH11231.

REFERENCES

1. Gutierrez T, Green DH, Nichols PD, Whitman WB, Semple KT, Aitken MD. 2013. *Polycyclovorans algicola* gen. nov., sp. nov., an aromatic-hydrocarbon-degrading marine bacterium found associated with laboratory cultures of marine phytoplankton. *Appl Environ Microbiol* 79: 205–214. <http://dx.doi.org/10.1128/AEM.02833-12>.
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, *Nephtunomonas 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. 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>.
7. 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>.
8. 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>.
9. 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>.
10. 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>.