AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements

Draft Genome Assembly of a Filamentous Euendolithic (True Boring) Cyanobacterium, *Mastigocoleus testarum* Strain BC008

Brandon S. Guida, Ferran Garcia-Pichel

Arizona State University, School of Life Sciences, Tempe, Arizona, USA

***Mastigocoleus testarum* strain BC008 is a model organism used to study marine photoautotrophic carbonate dissolution. It is a multicellular, filamentous, diazotrophic, euendolithic cyanobacterium ubiquitously found in marine benthic environments. We present an accurate draft genome assembly of 172 contigs spanning 12,700,239 bp with 9,131 annotated genes with an average G+C% of 37.3.**

Received 12 November 2015 **Accepted** 8 December 2015 **Published** 28 January 2016

Citation Guida BS, Garcia-Pichel F. 2016. Draft genome assembly of a filamentous euendolithic (true boring) cyanobacterium, *Mastigocoleus testarum* strain BC008. Genome Announc 4(1):e01574-15. doi:10.1128/genomeA.01574-15.

Copyright © 2016 Guida and Garcia-Pichel. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 Unported license.

Address correspondence to Ferran Garcia-Pichel, ferran@asu.edu.

Some microorganisms, known as euendoliths, have the remarkable capacity to excavate and grow into solid mineral phases, most typically carbonate minerals, often in the face of imposing physiological hurdles (1, 2). Cyanobacterial euendoliths are ubiquitous constituents of exposed carbonate outcrops, terrestrial and aquatic. In the marine environment they can represent important primary producers in carbonate-dominated systems (3). Euendolithic cyanobacteria play several significant roles in the dissolution of coral skeletons (4), and the erosion of coastal and terrestrial carbonate outcrops (5, 6). Their activity also impacts the viability of natural and farmed bivalve populations (7). Despite their ecological importance, currently no accurate genomic information is available for any euendolith.

Mastigocoleus testarum is a filamentous, true-branching, diazotrophic, morphologically and developmentally complex, cyanobacterial species, described as a euendolith in 1886 (8), an important pioneer of endolithic communities (9), and a pest for bivalve fisheries (10). A type strain, BC008, isolated from a shell fragment (11), represents the only working physiological model organism for the study of cyanobacterial euendoliths (1). While we have gained important insights into the mechanistic underpinning of carbonate dissolution using BC008, we are hampered by a lack of genetic information, which would enable comparative genomics, transcriptomics, and perhaps even the development of a genetic manipulation system and advances in its management in aquaculture.

M. testarum strain BC008 was grown as a unicyanobacterial, but not axenic, culture, in PES-30 medium (12) under a 16/8 h light/dark diel cycle at room temperature. Genomic DNA was isolated as previously described (13) and two separate Illumina MiSEQ paired-end (PE) libraries were created. One, 2 × 150 (lib1), was generated by the Joint Genome Institute and another, 2 × 300 (lib2), by the Translational Genomics Institute. The libraries revealed the presence of three distinct, divergent 16S rRNA sequences. In addition to the dominant 16S ribosomal sequence for *M. testarum*, the cultures contained 2 contaminants, one in the *Rhizobiales* and one related to *Hyphomonas* (99.7% identity). The genomic data set was thus treated as a metagenome. After quality filtering, both libraries combined contained over 7.2 Gbp of sequence. Each library underwent independent assembly using the iMETAMOS (14) pipeline of METAMOS (15). The best assembly was used for each library and the resultant contigs were assembled together using Geneious (16). Because both contaminants had a high G+C content genome (>50%), all contigs above that mark were removed. Contigs longer than 3 kbp were scaffolded using SSPACE (17), resulting in a final draft assembly of 172 scaffolds (>5 kbp, N_{50}:145,830; mean length: 65,170) representing 12,700,239 bp of genomic DNA with an average G+C% of 37.3, which is among the largest bacterial genomes reported. Gene prediction and annotation was done using the NCBI prokaryotic genome annotation pipeline (18) resulting in 9,131 genes, 73 tRNA loci, and 6 rRNA loci (two complete rRNA operons). Secondary metabolite gene cluster prediction was done using the antiSMASH server (19) and resulted in 22 predicted gene clusters indicating the potential biosynthesis of shinorine, nostopeptolide, hectochlorin, and staurosporine, among others.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number LMTZ00000000. The version described in this paper is version LMTZ01000000.

ACKNOWLEDGMENT

This work was primarily supported by NSF grant 1224939, "Intracellular metal pumping in microbial excavation by microbes."

FUNDING INFORMATION

National Science Foundation (NSF) provided funding to Ferran Garcia-Pichel under grant number 1224939.

REFERENCES

1. **Garcia-Pichel F, Ramirez-Reinat E, Gao Q.** 2010. Microbial excavation of solid carbonates powered by P-type ATPase-mediated transcellular Ca^{2+} transport. Proc Natl Acad Sci USA **107:**21749–21754. http://dx.doi.org/10.1073/pnas.1011884108.

2. **Garcia-Pichel F**. 2006. Plausible mechanisms for the boring on carbonates by microbial phototrophs. Sediment Geol **185:**205–213.
3. **Tribollet A, Langdon C, Golubic S, Atkinson M**. 2006. Endolithic microflora are major primary producers in dead carbonate substrates of Hawaiian coral reefs. J Phycol **42:**292–303. http://dx.doi.org/10.1111/j.1529-8817.2006.00198.x.
4. **Aline T**. 2008. Dissolution of dead corals by euendolithic microorganisms across the northern Great Barrier Reef (Australia). Microb Ecol **55:** 569–580. http://dx.doi.org/10.1007/s00248-007-9302-6.
5. **Trudgill ST, Smart PL, Friederich H, Crabtree RW**. 1987. Bioerosion of intertidal limestone, Co. Clare, Eire—1: *Paracentrotus lividus*. Mar Geol **74:**85–98.
6. **Shachak M, Jones CG, Granot Y**. 1987. Herbivory in rocks and the weathering of a desert. Mar Geol **236:**1098–1099. http://dx.doi.org/10.1126/science.236.4805.1098.
7. **Kosnik M, Zardi GI, Nicastro KR, McQuaid CD, Gektidis M**. 2009. Effects of endolithic parasitism on invasive and indigenous mussels in a variable physical environment. PLoS One **4:**e6560.
8. **Lagerheim G**. 1886. Note sur le mastigocoleus, nouveau genre des algues marines de l' ordre des phycochromacées. Notarisia **1:**65–69.
9. **Le Campion-Alsumard T, Golubic S, Hutchings P**. 1995. Microbial endoliths in skeletons of live and dead corals: *Porites lobata* (Moorea, French Polynesia). Mar Ecol Prog Ser **117:**149–157. http://dx.doi.org/10.3354/meps117149.
10. **Webb SC, Korrûbel JL**. 1994. Shell weakening in marine mytilids attributable to blue-green alga *Mastigocoleus* sp. (*Nostochopsidaceae*). J Shellfish Res **13:**11–17.
11. **Chacón E, Berrendero E, Garcia-Pichel F**. 2006. Biogeological signatures of microboring cyanobacterial communities in marine carbonates from Cabo Rojo, Puerto Rico. Sedimentary Geol **185:**215–228.
12. **Provasoli L**. 1966, Media and prospects for the cultivation of marine algae, p 63–75. Japanese Society of Plant Physiology, Kyoto, Japan.
13. **Ramírez-Reinat EL, Garcia-Pichel F**. 2012. Characterization of a marine cyanobacterium that bores into carbonates and the redescription of the genus *Mastigocoleus*. J Phycol **48:**740–749. http://dx.doi.org/10.1111/j.1529-8817.2012.01157.x.
14. **Koren S, Treangen TJ, Hill CM, Pop M, Phillippy AM**. 2014. Automated ensemble assembly and validation of microbial genomes. BMC Bioinformatics **15:**126. http://dx.doi.org/10.1186/1471-2105-15-126.
15. **Treangen TJ, Koren S, Astrovskaya I, Sommer D, Liu B, Pop M**. 2011. MetAMOS: A metagenomic assembly and analysis pipeline for AMOS. Genome Biol **12:**25. http://dx.doi.org/10.1186/gb-2011-12-s1-p25.
16. **Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A**. 2012. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics **28:**1647–1649. http://dx.doi.org/10.1093/bioinformatics/bts199.
17. **Boetzer M, Pirovano W**. 2014. SSPACE-LongRead: scaffolding bacterial draft genomes using long read sequence information. BMC Bioinformatics **15:**211. http://dx.doi.org/10.1186/1471-2105-15-211.
18. **Angiuoli SV, Gussman A, Klimke W, Cochrane G, Field D, Garrity GM, Kodira CD, Kyrpides N, Madupu R, Markowitz V, Tatusova T, Thomson N, White O**. 2008. Toward an online repository of standard operating procedures (SOPs) for (meta)genomic annotation. OMICS **12:**137–141. http://dx.doi.org/10.1089/omi.2008.0017.
19. **Medema MH, Blin K, Cimermancic P, de Jager V, Zakrzewski P, Fischbach MA, Weber T, Takano E, Breitling R**. 2011. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. Nucleic Acids Res **39:**W339–W346. http://dx.doi.org/10.1093/nar/gkr466.