Draft Genome Sequence of "*Candidatus Synechococcus spongiarum*" m9, Binned from a Metagenome of South China Sea Sponge *Theonella swinhoei*

Fang Liu, Jinlong Li, Zhiyong Li

Marine Biotechnology Laboratory, State Key Laboratory of Microbial Metabolism and School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, People's Republic of China

ABSTRACT "*Candidatus* Synechococcus spongiarum" represents the widespread cyanobacterial symbionts found in marine sponges with relatively high genomic variability and likely important ecological roles. We present here the draft genome sequence of "*Candidatus* Synechococcus spongiarum" m9, which was assembled from a metagenome of *Theonella swinhoei* sampled in the South China Sea.

Members of "*Candidatus* Synechococcus spongiarum" are widespread in sponge symbiotic microbial communities (1). These cyanobacterial symbionts are uncultivated and may play important roles in primary production, UV protection, and chemical defense (2, 3). Application of metagenome binning and single-cell sequencing has revealed four "*Ca.* Synechococcus spongiarum" draft genomes from the sponges sampled in the Red Sea and the Mediterranean Sea (4, 5). Nonetheless, these four genomes varied in genome size, gene content, and immune system features, although their 16S rRNA gene sequences were highly similar (~99%) (5). Hence, it is necessary to obtain more genomes from different locations/hosts for the better understanding of the genomic diversity and adaptation mechanisms of "*Ca.* Synechococcus spongiarum." None of the available "*Ca.* Synechococcus spongiarum" genomes are from the Pacific Ocean. Here, we describe the "*Ca.* Synechococcus spongiarum" m9 draft genome sequence recovered from the metagenome of a South China Sea sponge, *Theonella swinhoei*.

Theonella swinhoei was collected near Yongxing Island (112°20′E, 16°50′N) in the South China Sea. DNA extraction, Illumina sequencing, and metagenome assembling were carried out as described by Liu et al. (6). VizBin was used to extract the metagenomic bins from the metagenome contigs, using default settings (7). Contigs shorter than 2,000 bp were removed from each bin. PhyloSift (8) and JSpeciesWS (9) were used to determine the phylogenetic affiliation of each bin. The completeness, contamination rate, and average amino acid index (AAI) were calculated by enveomics-GUI (https://github.com/lmrodriguezr/enveomics-gui), according to Liu et al. (6). The RASTtk pipeline (10) and PATRIC (11) were used for genome annotation and comparative analysis. Additionally, WebMGA (12) was used for COG annotation.

The above-described process resulted in a cyanobacterial metagenome bin, namely, "*Ca.* Synechococcus spongiarum" m9. This genome consists of 179 contigs with a genome size of 1.4 Mb (N_{50}, 10.7 kb) and a G+C content of 61.9%. enveomics-GUI indicated a genome completeness of 93%, with likely no contamination. Phylogenetically, it is closely related to "*Ca.* Synechococcus spongiarum" SP3 from the *T. swinhoei* sampled from the Red Sea (5), as suggested by a higher AAI (86.87%) than the AAI values between other public "*Ca.* Synechococcus spongiarum" draft genomes (approx. 84 to 85%). The m9 genome contains 1,659 coding sequences (CDSs) (585 hypothetical

proteins) and 33 tRNAs. It is smaller than any known "*Ca.* Synechococcus spongiarum" genome (5). This cyanobacterial symbiont has a metabolism repertoire similar to that of the other "*Ca.* Synechococcus spongiarum" strains, based on PATRIC analysis. Seventy-nine percent of the previously defined "*Ca.* Synechococcus spongiarum" "core COGs" (5)" exist in the m9 genome. Like its relatives, the m9 genome is short of antioxidant enzymes and small peptides of photosystem II, which are regarded to be adaptive traits to the symbiotic lifestyle (5). Nonetheless, genes of biosynthesis of dTDP-L-rhamnose (O antigen), the absence of which was regarded to be a signature of "*Ca.* Synechococcus spongiarum" (5), were found in m9. This observation underlines the variability of "*Ca.* Synechococcus spongiarum" genomes and suggests the diverse mechanisms in sponge-cyanobacterium recognition.

Accession number(s). The "*Ca.* Synechococcus spongiarum" m9 draft genome has been deposited at ENA with the contig accession range FITM01000001 to FITM01000179.

REFERENCES

1. Erwin PM, López-Legentil S, Turon X. 2012. Ultrastructure, molecular phylogenetics, and chlorophyll a content of novel cyanobacterial symbionts in temperate sponges. Microb Ecol 64:771–783. https://doi.org/10.1007/s00248-012-0047-5.
2. Taylor MW, Radax R, Steger D, Wagner M. 2007. Sponge-associated microorganisms: evolution, ecology, and biotechnological potential. Microbiol Mol Biol Rev 71:295–347. https://doi.org/10.1128/MMBR.00040-06.
3. Webster NS, Taylor MW. 2012. Marine sponges and their microbial symbionts: love and other relationships. Environ Microbiol 14:335–346. https://doi.org/10.1111/j.1462-2920.2011.02640.x.
4. Gao ZM, Wang Y, Tian RM, Wong YH, Batang ZB, Al-Suwailem AM, Bajic VB, Qian PY. 2014. Symbiotic adaptation drives genome streamlining of the cyanobacterial sponge symbiont "Candidatus Synechococcus spongiarum." mBio 5:e00079-14. https://doi.org/10.1128/mBio.00079-14.
5. Burgsdorf I, Slaby BM, Handley KM, Haber M, Blom J, Marshall CW, Gilbert JA, Hentschel U, Steindler L. 2015. Lifestyle evolution in cyanobacterial symbionts of sponges. mBio 6:e00391-15. https://doi.org/10.1128/mBio.00391-15.
6. Liu F, Li J, Feng G, Li Z. 2016. New genomic insights into "*Entotheonella*" symbionts in *Theonella swinhoei*: mixotrophy, anaerobic adaptation, resilience, and interaction. Front Microbiol 7:1333 https://doi.org/10.3389/fmicb.2016.01333.
7. Laczny CC, Sternal T, Plugaru V, Gawron P, Atashpendar A, Margossian HH, Coronado S, der Maaten Lv, Vlassis N, Wilmes P. 2015. VizBin—an application for reference-independent visualization and human-augmented binning of metagenomic data. Microbiome 3:1. https://doi.org/10.1186/s40168-014-0066-1.
8. Darling AE, Jospin G, Lowe E, Matsen FAIV, Bik HM, Eisen JA. 2014. PhyloSift: phylogenetic analysis of genomes and metagenomes. PeerJ 2:e243. https://doi.org/10.7717/peerj.243.
9. Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J. 2016. JSpeciesWS: a Web server for prokaryotic species circumscription based on pairwise genome comparison. Bioinformatics 32:929–931. https://doi.org/10.1093/bioinformatics/btv681.
10. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA, Stevens R, Vonstein V, Wattam AR, Xia F. 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep 5:8365 https://doi.org/10.1038/srep08365.
11. Wattam AR, Abraham D, Dalay O, Disz TL, Driscoll T, Gabbard JL, Gillespie JJ, Gough R, Hix D, Kenyon R, MacHi D, Mao C, Nordberg EK, Olson R, Overbeek R, Pusch GD, Shukla M, Schulman J, Stevens RL, Sullivan DE, Vonstein V, Warren A, Will R, Wilson MJC, Yoo HS, Zhang C, Zhang Y, Sobral BW. 2014. PATRIC, the bacterial bioinformatics database and analysis resource. Nucleic Acids Res 42:D581–D591. https://doi.org/10.1093/nar/gkt1099.
12. Wu S, Zhu Z, Fu L, Niu B, Li W. 2011. WebMGA: a customizable Web server for fast metagenomic sequence analysis. BMC Genomics 12:444. https://doi.org/10.1186/1471-2164-12-444.