# Complete Genome Sequence of *Thalassotalea euphylliae* Strain H2

Stephen Summers,ᵃ Marnie L. Freckelton,ᵇ Brian T. Nedved,ᵇ Scott A. Rice,ᵃ,ᶜ Michael G. Hadfieldᵇ

ᵃSingapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University, Singapore
ᵇKewalo Marine Laboratory, University of Hawai'i, Honolulu, Hawaii, USA
ᶜThe School of Biological Sciences, Nanyang Technological University, Singapore

**ABSTRACT** A bacterial isolate of *Thalassotalea euphylliae* H2 was collected from the coral *Montipora capitata*. MinION long reads were employed for scaffolding and complemented with short-read MiSeq sequences to permit complete genome assembly. The genome is approximately 4.36 Mb long, with 3,669 protein-coding genes, 92 tRNAs, and 21 rRNAs.

The bacterium *Thalassotalea euphylliae* is a Gram-negative rod-shaped bacterium isolated from the coral *Euphylliae glabrescens* (1). This bacterium is motile, having a single polar flagellum, and it forms a pale-yellow colony. Other members of the species *Thalassotalea euphylliae* have also been isolated and their genomes reported (2). This bacterium has been associated with a range of marine environments, such as oyster tissues (3), tidal flats (4), marine sediments (5), and seawater (6).

Here, we present the genome of *Thalassotalea euphylliae* H2, isolated from the mucus of the coral *Montipora capitata*. The coral was sampled on the reef adjacent to the Hawaii Institute of Marine Biology (University of Hawai'i) in Kane'ohe Bay, O'ahu, Hawaii, in 2010 (7). The bacterium was isolated from the surface of coral branches by streak plating on half-strength seawater-tryptone agar (8). The culture is retained at −80°C at the Kewalo Marine Laboratory (University of Hawai'i, Honolulu, HI).

The isolate was grown overnight at 28°C in a half-strength seawater-tryptone broth (8), and the cells were collected by centrifugation at $4,000 \times g$ for 5 min. Total nucleic acids were extracted from the cell pellets using a xanthogenate-based extraction technique (9). Following extraction, 300-bp paired-end sequencing was carried out on the Illumina MiSeq platform after library preparation using a TruSeq Nano DNA kit, which yielded 2,534,742 reads. Illumina MiSeq sequencing adapters were removed using the Trimmomatic package (version 0.36), and quality control (QC) was performed using FastQC (version 0.11.7). Assembly was assisted by the use of long-read sequencing as a scaffold. This was performed on a Nanopore MinION genome sequencer using a rapid sequencer kit (SQK-RAD004) and a R9.4 flow cell. This resulted in 536,912 reads of up to 131,642 bp in length (median, 5,874 bp).

The long reads were base called using the Albacore package version 2.3.1, as supplied by Nanopore, using default settings, and adapters were removed using Porechop version 0.2.3 (verbosity level 3). The *de novo* assembly was conducted using the Unicycler package version 0.4.4 (10) using long reads as a scaffold to enhance MiSeq assembly of the genome. Polishing and rotation were performed using the default settings in Racon version 1.3.1 (11).

The final polished circular chromosomal contig was 4,357,276 bp (174-fold coverage, 43.0% GC content). Species identification was achieved by BLASTn comparison of the NCBI 16S rRNA gene database and showed a 98% match to *T. euphylliae* strain Eup-16 (NCBI RefSeq accession number NR_153727).

Genome annotation was completed with the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) version 4.2 (12). A total of 3,808 genes were identified, consisting of 3,669 protein-coding genes and 115 RNAs (92 tRNAs, 4 noncoding RNAs, 7 5S rRNA subunits, and 6 copies each of the 16S and 23S rRNAs). Of these, 31 genes were identified as related to membrane function and maintenance. Using the functional annotation services provided by the Rapid Annotations using Subsystems Technology (RAST [13]) servers, we identified 3,897 DNA-coding sequences within 151 subsystems.

**Data availability.** The genome sequence of *T. euphylliae* H2 has been deposited in GenBank under accession number QUOV00000000. Raw sequencing reads are available in the NCBI Sequence Read Archive (accession numbers SRR8069230 and SRR8069231).

## ACKNOWLEDGMENTS

## REFERENCES

1. Sheu S-Y, Liu L-P, Tang S-L, Chen W-M. 2016. *Thalassotalea euphylliae* sp. nov., isolated from the torch coral *Euphyllia glabrescens*. Int J Syst Evol Microbiol 66:5039–5045. https://doi.org/10.1099/ijsem.0.001466.

2. Summers S, Freckelton ML, Nedved BT, Rice SA, Hadfield MG. 2018. Full-genome sequence of *Thalassotalea euphylliae* H1, isolated from a *Montipora capitata* coral located in Hawai'i. Microbiol Resour Announc 7:e01244-18. https://doi.org/10.1128/MRA.01244-18.

3. Choi S, Kim E, Shin S-K, Yi H. 2017. *Thalassotalea crassostreae* sp. nov., isolated from Pacific oyster. Int J Syst Evol Microbiol 67:2195–2198. https://doi.org/10.1099/ijsem.0.001923.

4. Park S, Choi J, Won S-M, Yoon J-H. 2018. *Thalassotalea insulae* sp. nov., isolated from tidal flat sediment. Int J Syst Evol Microbiol 68:1321–1326. https://doi.org/10.1099/ijsem.0.002671.

5. Kang H, Kim H, Nam I-Y, Joung Y, Jang TY, Joh K. 2017. *Thalassotalea litorea* sp. nov., isolated from seashore sand. Int J Syst Evol Microbiol 67:2268–2273. https://doi.org/10.1099/ijsem.0.001938.

6. Wang Y, Liu T, Ming H, Sun P, Cao C, Guo M, Du J, Zhou C, Zhu W. 2018. *Thalassotalea atypica* sp. nov., isolated from seawater, and emended description of *Thalassotalea eurytherma*. Int J Syst Evol Microbiol 68:271–276. https://doi.org/10.1099/ijsem.0.002495.

7. Tran C, Hadfield MG. 2011. Larvae of *Pocillopora damicornis* (Anthozoa) settle and metamorphose in response to surface-biofilm bacteria. Mar Ecol Prog Ser 433:85–96. https://doi.org/10.3354/meps09192.

8. Boettcher K, Ruby E. 1990. Depressed light emission by symbiotic *Vibrio fischeri* of the sepiolid squid *Euprymna scolopes*. J Bacteriol 172:3701–3706. https://doi.org/10.1128/jb.172.7.3701-3706.1990.

9. Tillett D, Neilan BA. 2000. Xanthogenate nucleic acid isolation from cultured and environmental cyanobacteria. J Phycol 36:251–258. https://doi.org/10.1046/j.1529-8817.2000.99079.x.

10. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595.

11. Vaser R, Sovic I, Nagarajan N, Sikic M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res 27:737–746. https://doi.org/10.1101/gr.214270.116.

12. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. Nucleic Acids Res 44:6614–6624. https://doi.org/10.1093/nar/gkw569.

13. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res 42:D206–D214. https://doi.org/10.1093/nar/gkt1226.