AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

| Environmental Microbiology | Announcement

Complete genome sequence of a potential new species *Vibrio* sp. NTOU-M3 isolated from hard clam, *Meretrix taiwanica*, in Taiwan

Che-Chun Chen,[1,2] Wei-Hsiang Lin,[3] Te-Hua Hsu,[4,5] Ying-Ning Ho[1,2,3,5]

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT *Vibrio* sp. NTOU-*M3* is a potential new bacterium isolated from hard clam (*Meretrix taiwanica*) in the estuarine region of Taiwan. The complete sequences obtained using Oxford Nanopore Technologies and Illumina sequencing consist of a 3,272,438-bp large circular chromosome and a 1,584,497-bp small circular chromosome.

KEYWORDS *Vibrio*, hard clam, *Meretrix taiwanica*, Oxford Nanopore Technologies

The genus *Vibrio* consists of a group of marine bacteria that infect several genera of marine shellfish (1). The novel *Vibrio* sp. NTOU-M3 was isolated from hard clam (*Meretrix taiwanica*) obtained from Heping Island Fish Market, Keelung City, Taiwan, and sourced from the estuarine region of Taiwan. First, 0.1 g tissue was collected and mixed with 2 mL ddH$_2$O, and then 100 µL mixture was added with 1/10 marine broth (1 mL). Finally, it was dropped on 1/10 marine broth agar plates (HiMedia Laboratorie, India). The plates were incubated at 37°C for 12 hours. A single colony purified by three rounds of streaking was selected for identification. Genomic DNA for both Illumina and Nanopore sequencing was extracted using the Nanobind CBB Kit (PacBio, USA). The DNA quantity was determined using the Qubit dsDNA BR assay kit (Thermo Scientific, USA). The 16S rRNA gene sequence was amplified by primers 27F and 1492R (2). Sanger sequencing yielded the nearly complete 16S rRNA gene sequence (PP981031). This 16S rRNA sequence was compared by NCBI BLASTn and limit to sequences from the type material. The most similar species is *Vibrio harveyi* NBRC 15634 (98.30%, NR_113784). The optimal identity thresholds for potential new species were ~98.7% (3).

The Illumina sequencing library was generated using the TruSeq Nano DNA Library Prep Kit (Illumina), according to the manufacturer's instructions. Genomic DNA was fragmented by sonication at 350 bp and then purified using Sample Purification Beads (Illumina). The NovaSeq X Plus platform was used by Genomics BioSci & Tech to generate 150-bp paired-end reads. Raw Illumina reads (total reads: 10,106,288) were trimmed using Trimmomatic v0.39 (4) and quality-controlled using fastp v0.23.4 (5). Long-read sequencing library was prepared using a the Nanopore genomic DNA ligation kit (SQK-LSK110) and sequenced using a Flongle flow cell R9.4.1 (FLO-FLG001) for 96 hours on the MinION platform. Nanopore data were basecalled using Dorado v0.6.0 (https://github.com/nanoporetech/dorado) and filtered by Nanofilt v2.8.0 (6) to generate Q12 data (raw reads: 50,476; N50: 32,611). Assembly was performed using Nanophase v0.2.3 (7) with the hybrid method, resulting in two circular chromosomes (average coverage: 243×) and rotated to start at the dnaA gene using the Circlator v1.5.5-docker5 (8). The genome includes two chromosomes of 4,856,935 bp with a G + C content of 43.8%. The final genome was annotated using DFAST v1.6.0 (9), revealing 4,279 CDSs, 28 rRNA genes, and 105 tRNA genes. The genome assembly had a completeness of 100% and contamination level of 0%, as confirmed by using CheckM v1.2.2 (10). The novel *Vibrio* sp. shows low ANI value (75.79%) and the dDDH (24%) to the closest type strain, *Vibrio harveyi* NBRC 15634 via CJ Bioscience's online ANI calculator v0.91 (11)

Editor Frank J. Stewart, Montana State University, Bozeman, Montana, USA

Address correspondence to Ying-Ning Ho, ynho@email.ntou.edu.tw.

The authors declare no conflict of interest.

See the funding table on p. 2.

Received 25 July 2024
Accepted 12 September 2024
Published 27 September 2024

Copyright © 2024 Chen et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

and TYGS v391 (12). In addition, antiSMASH 7.0 (13) analysis identified four secondary metabolite biosynthetic gene clusters, namely, ectoine, hserlactone, NRP-metallophore, and arylpolyene, in the large chromosome and three BGCs including a betalactone and two RiPP-like in the small chromosome of the *Vibrio* sp. NTOU-M3. The contributions of this complete genome sequence encourage further research into the relationship between *Vibrio* and hard clam *M. taiwanica*.

ACKNOWLEDGMENTS

This study was supported by the National Science and Technology Council of Taiwan (MOST 111–2628 M-019–001-MY3).

AUTHOR AFFILIATIONS

[1]Doctoral Degree Program in Marine Biotechnology, National Taiwan Ocean University, Keelung, Taiwan
[2]Taiwan Oceans Genome Center, National Taiwan Ocean University, Keelung, Taiwan
[3]Institute of Marine Biology, National Taiwan Ocean University, Keelung, Taiwan
[4]Department of Aquaculture, National Taiwan Ocean University, Keelung, Taiwan
[5]Center of Excellence for the Oceans, National Taiwan Ocean University, Keelung, Taiwan

AUTHOR ORCIDs

Ying-Ning Ho http://orcid.org/0000-0003-0943-1416

FUNDING

Funder	Grant(s)	Author(s)
National Science and Technology Council (NSTC)	MOST 111-2628-M-019-001-MY3	Ying-Ning Ho

AUTHOR CONTRIBUTIONS

Che-Chun Chen, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Writing – original draft | Wei-Hsiang Lin, Data curation, Investigation, Methodology, Validation | Te-Hua Hsu, Conceptualization, Data curation, Resources, Supervision | Ying-Ning Ho, Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Project administration, Software, Supervision, Writing – review and editing

DATA AVAILABILITY

Default parameters were used except where otherwise noted. This genome project has been deposited in the National Center for Biotechnology Information (NCBI) with Bioproject accession number PRJNA1136137 and BioSample accession number SAMN42503066. The raw reads are available at the NCBI Sequence Read Archive under accession number SRX25366312 to SRX25366313. The genome sequence can be found in GenBank under accession number CP162100 to CP162101.

REFERENCES

1. Panicker G, Call DR, Krug MJ, Bej AK. 2004. Detection of pathogenic *Vibrio* spp. in shellfish by using multiplex PCR and DNA microarrays. Appl Environ Microbiol 70:7436–7444. https://doi.org/10.1128/AEM.70.12.7436-7444.2004
2. Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner FO. 2013. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res 41:e1. https://doi.org/10.1093/nar/gks808
3. Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, Schleifer K-H, Whitman WB, Euzéby J, Amann R, Rosselló-Móra R. 2014. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nat Rev Microbiol 12:635–645. https://doi.org/10.1038/nrmicro3330
4. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170
5. Chen S. 2023. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. Imeta 2:e107. https://doi.org/10.1002/imt2.107

6. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. Bioinformatics 34:2666–2669. https://doi.org/10.1093/bioinformatics/bty149
7. Liu L, Yang Y, Deng Y, Zhang T. 2022. Nanopore long-read-only metagenomics enables complete and high-quality genome reconstruction from mock and complex metagenomes. Microbiome 10:209. https://doi.org/10.1186/s40168-022-01415-8
8. Hunt M, Silva ND, Otto TD, Parkhill J, Keane JA, Harris SR. 2015. Circlator: automated circularization of genome assemblies using long sequencing reads. Genome Biol 16:294. https://doi.org/10.1186/s13059-015-0849-0
9. Tanizawa Y, Fujisawa T, Nakamura Y. 2018. DFAST: a flexible prokaryotic genome annotation pipeline for faster genome publication. Bioinformatics 34:1037–1039. https://doi.org/10.1093/bioinformatics/btx713
10. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055. https://doi.org/10.1101/gr.186072.114
11. Yoon S-H, Ha S-M, Lim J, Kwon S, Chun J. 2017. A large-scale evaluation of algorithms to calculate average nucleotide identity. Antonie Van Leeuwenhoek 110:1281–1286. https://doi.org/10.1007/s10482-017-0844-4
12. Meier-Kolthoff JP, Göker M. 2019. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. Nat Commun 10:2182. https://doi.org/10.1038/s41467-019-10210-3
13. Blin K, Shaw S, Augustijn HE, Reitz ZL, Biermann F, Alanjary M, Fetter A, Terlouw BR, Metcalf WW, Helfrich EJN, van Wezel GP, Medema MH, Weber T. 2023. antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. Nucleic Acids Res 51:W46–W50. https://doi.org/10.1093/nar/gkad344