



A *Mollicutes* Metagenome-Assembled Genome from the Gut of the Pteropod *Limacina rangii*

Zachary T. Pimentel,^a  Patricia S. Thibodeau,^{b*} Bongkeun Song,^b  Ying Zhang^a

^aDepartment of Cell and Molecular Biology, College of Environment and Life Sciences, University of Rhode Island, Kingston, Rhode Island, USA

^bDepartment of Biological Sciences, Virginia Institute of Marine Science, William & Mary, Gloucester Point, Virginia, USA

ABSTRACT A nearly complete genome of an uncultured *Mollicutes* sp. was obtained from the metagenome of the gut of *Limacina rangii* (open-ocean snail), an important grazer and prey for higher trophic animals along the rapidly warming region of the western Antarctic Peninsula.

A recent metabarcoding study of *Limacina rangii*, a dominant grazer among zooplankton and an important prey along the western Antarctic Peninsula (WAP), revealed that *Mollicutes* bacteria are a cosmopolitan and dominant component of the gut microbiome (1, 2). Here, we report a nearly complete metagenome-assembled genome (MAG) of the class *Mollicutes* from the *L. rangii* gut obtained along the WAP, contributing to the growing number of genomic resources available for host-associated marine *Mollicutes* (3–8).

L. rangii organisms were collected at selected stations along the WAP (600.200, 300.200, and 100.040) in January 2017 (9). Gut samples were dissected onboard and immediately frozen at -80°C . Total DNA was extracted from three gut samples per station (9), using the Qiagen DNeasy blood and tissue kit, and then pooled to create a single extract for sequencing. Libraries were prepared with 30 to 50 ng DNA per sample using the Nextera DNA sample preparation kit (Illumina).

Sequencing was performed at a read length of 2×150 bp by MR DNA (Molecular Research LP) on the Illumina HiSeq 2500 platform. Raw read quality was visualized with FASTQC version 0.111.14 (10). A total of 31,260,228 read pairs across three samples underwent quality filtering with Trimmomatic version 0.38 (11) with the following parameters: minimum length of 90 bp, 4-bp sliding window with an average quality score of 15, and leading/trailing bases with quality scores less than 3 were removed. Quality-filtered reads were coassembled using default parameters with MEGAHIT version 1.1.1 (12). The coassembly was indexed with bowtie2 version 2.2.9, and the quality-filtered reads from each sample were mapped to the coassembly (13). SAM files were converted to BAM format and ordered using SAMtools release 1.5 (14). MetaBAT2 version 2.12.1 (15) was used to bin the coassembled contigs with default parameters. The *lineage_wf* function in CheckM version 1.0.5 (16) estimated completeness and contamination. Gene prediction and annotation were completed with PGAP (17). Conserved single-copy genes (CSCGs) were identified through analysis of bidirectional best BLAST hits between the MAG, *Mycoplasma* reference genomes, and outgroup *Firmicutes* genomes. Each CSCG cluster was aligned with MUSCLE version 3.8.3, and a phylogenetic reconstruction was performed with RAxML version 8.2.10 using the JTT substitution model and the GAMMA model of rate heterogeneity as previously described (3).

In total, 637,502 contigs were coassembled, but only one MAG, taxonomically assigned to the class *Mollicutes*, was obtained with high estimated completeness. The mean depth of coverage was highest at the most northern site (600.200) at $22.9\times$, while it was less than $2\times$ at the other sites (300.200 and 100.040). The *Mollicutes* MAG contained 85 contigs and a

Editor Frank J. Stewart, Montana State University

Copyright © 2022 Pimentel et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Ying Zhang, yingzhang@uri.edu, or Patricia S. Thibodeau, thibodeau@uri.edu.

*Present address: Patricia S. Thibodeau, Graduate School of Oceanography, University of Rhode Island, Narragansett, Rhode Island, USA.

The authors declare no conflict of interest.

Received 29 July 2022

Accepted 3 November 2022

Published 16 November 2022

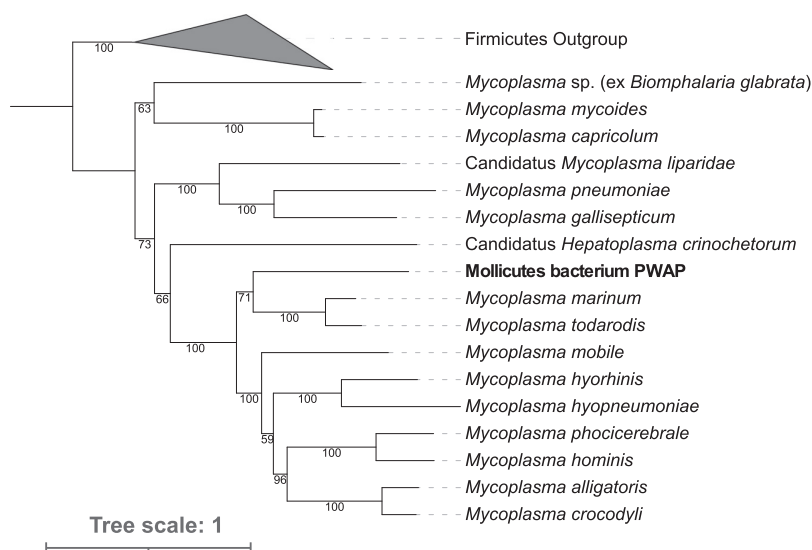


FIG 1 Maximum-likelihood phylogeny based on conserved single-copy genes between the *Mollicutes* MAG from this study (PWAP), reference *Mycoplasma* genomes, and four genomes from the *Firmicutes* used as an outgroup. One hundred iterations were used to compute bootstrap values. The GenBank accession numbers for the genomes in the phylogenetic reconstruction are included in Table 1.

genome size of 0.55 Mb, with an N_{50} of 7,347 bp, estimated completeness of 88.16%, contamination and strain heterogeneity of 0%, and GC content of 25.1%. The *Mollicutes* MAG included 527 protein-coding genes and contained multiple rRNA genes (one each of 16S, 23S, and 5S). The MAG was most closely related to *Mycoplasma marinum* and *Mycoplasma todarodis*, which were isolated from an octopus and squid, respectively, based on a phylogenetic reconstruction from 63 CSCGs (Fig. 1; Table 1).

Data availability. The raw reads were deposited in the NCBI SRA database with accession numbers [SRR12228976](https://www.ncbi.nlm.nih.gov/sra/SRR12228976), [SRR12228977](https://www.ncbi.nlm.nih.gov/sra/SRR12228977), and [SRR12228978](https://www.ncbi.nlm.nih.gov/sra/SRR12228978), and the MAG assembly is available under [ASM1966174v1](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA646234), all attached to BioProject [PRJNA646234](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA646234).

TABLE 1 NCBI accession numbers for bacterial genomes included in the phylogenetic reconstruction

Organism	NCBI RefSeq accession no.
<i>Mycoplasma pneumoniae</i> M129	GCF_000027345.1
<i>Mycoplasma gallisepticum</i>	GCF_001676495.1
"Candidatus <i>Mycoplasma liparidae</i> "	GCA_009884515.1
<i>Mycoplasma</i> sp. (ex <i>Biomphalaria glabrata</i>)	GCF_001484045.1
<i>Mycoplasma hyopneumoniae</i>	GCF_002257505.1
"Candidatus <i>Hepatoplasma crinochetorum</i> " Av	GCF_000582535.1
<i>Mycoplasma mycoides</i> subsp. <i>capri</i>	GCF_900489525.1
<i>Mycoplasma alligatoris</i> A21JP2	GCF_000178375.1
<i>Mycoplasma capricolum</i> subsp. <i>capripneumoniae</i> 87001	GCF_000835085.1
<i>Mycoplasma crocodyli</i> MP145	GCF_000025845.1
<i>Mycoplasma hyorhinis</i>	GCF_001705605.1
<i>Mycoplasma hominis</i>	GCF_000759375.2
<i>Mycoplasma phocicerebrale</i>	GCF_003383595.3
<i>Mycoplasma todarodis</i>	GCF_004335995.1
<i>Mycoplasma mobile</i> 163K	GCF_000008365.1
<i>Mycoplasma marinum</i>	GCF_004335975.1
<i>Lactobacillus plantarum</i> WCFS1	GCF_000203855.3
<i>Listeria monocytogenes</i> EGD-e	GCF_000196035.1
<i>Enterococcus faecalis</i> V583	GCF_000007785.1
<i>Staphylococcus aureus</i> subsp. <i>aureus</i> NCTC 8325	GCF_000013425.1

ACKNOWLEDGMENTS

Funding was from the National Science Foundation Antarctic Organisms and Ecosystems Program (PLR 1440435), NSF Biological Oceanography (OCE 1321373 and 1737258), NSF OIA-1929078, and Rhode Island NSF EPSCoR grant OIA-165522.

Z.T.P. wrote the manuscript and analyzed sequence data. P.S.T. wrote the manuscript and provided sequence data. B.S. provided sequence data and revised the manuscript. Y.Z. provided analysis oversight and revised the manuscript.

REFERENCES

- Janssen AW, Bush SL, Bednaršek N. 2019. The shelled pteropods of the northeast Pacific Ocean (Mollusca: Heterobranchia, Pteropoda). *Zoosymposia* 13:305–346. <https://doi.org/10.11646/zoosymposia.13.1.22>.
- Thibodeau PS, Song B, Moreno CM, Steinberg DK. 2022. Feeding ecology and microbiome of the pteropod *Limacina helicina antarctica*. *Aquat Microb Ecol* 88:19–24. <https://doi.org/10.3354/ame01981>.
- Pimentel ZT, Dufault-Thompson K, Russo KT, Scro AK, Smolowitz RM, Gomez-Chiarri M, Zhang Y. 2021. Microbiome analysis reveals diversity and function of mollicutes associated with the Eastern oyster, *Crassostrea virginica*. *mSphere* 6:e00227-21. <https://doi.org/10.1128/mSphere.00227-21>.
- Rasmussen JA, Villumsen KR, Duchêne DA, Puetz LC, Delmont TO, Sveier H, Jørgensen L von G, Præbel K, Martin MD, Bojesen AM, Gilbert MTP, Kristiansen K, Limborg MT. 2021. Genome-resolved metagenomics suggests a mutualistic relationship between *Mycoplasma* and salmonid hosts. *Commun Biol* 4:579. <https://doi.org/10.1038/s42003-021-02105-1>.
- Ramírez AS, Vega-Orellana OM, Viver T, Poveda JB, Rosales RS, Poveda CG, Spersger J, Szostak MP, Caballero MJ, Ressel L, Bradbury JM, Mar Tavío M, Karthikeyan S, Amann R, Konstantinidis KT, Rossello-Mora R. 2019. First description of two moderately halophilic and psychrotolerant *Mycoplasma* species isolated from cephalopods and proposal of *Mycoplasma marinum* sp. nov. and *Mycoplasma todarodis* sp. nov. *Syst Appl Microbiol* 42:457–467. <https://doi.org/10.1016/j.syapm.2019.04.003>.
- Cheab B, Yang P, Kazlauskaitė R, Lindsay E, Heys C, Dwyer T, De Noia M, Schaal P, Sloan W, Ijaz UZ, Llewellyn MS. 2021. Genome erosion and evidence for an intracellular niche – exploring the biology of mycoplasmas in Atlantic salmon. *Aquaculture* 541:736772. <https://doi.org/10.1016/j.aquaculture.2021.736772>.
- He L-S, Zhang P-W, Huang J-M, Zhu F-C, Danchin A, Wang Y. 2018. The enigmatic genome of an obligate ancient *Spiroplasma* symbiont in a hadal holothurian. *Appl Environ Microbiol* 84:e01965-17. <https://doi.org/10.1128/AEM.01965-17>.
- Wang Y, Huang J-M, Wang S-L, Gao Z-M, Zhang A-Q, Danchin A, He L-S. 2016. Genomic characterization of symbiotic mycoplasmas from the stomach of deep-sea isopod *Bathynomus* sp. *Environ Microbiol* 18: 2646–2659. <https://doi.org/10.1111/1462-2920.13411>.
- Thibodeau PS, Steinberg DK, Stammerjohn SE, Hauri C. 2019. Environmental controls on pteropod biogeography along the Western Antarctic Peninsula. *Limnol Oceanogr* 64:S240–S256. <https://doi.org/10.1002/lno.11041>.
- Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom.
- 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>.
- Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31:1674–1676. <https://doi.org/10.1093/bioinformatics/btv033>.
- Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10:R25. <https://doi.org/10.1186/gb-2009-10-3-r25>.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
- Kang D, Li F, Kirton ES, Thomas A, Egan RS, An H, Wang Z. 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ Preprints* e27522v1. <https://doi.org/10.7717/peerj.7359>.
- 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>.
- 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>.