AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of the *Arcobacter marinus* Type Strain JCM 15502

William G. Miller,[a] Emma Yee,[a] Steven Huynh,[a] Craig T. Parker[a]

[a]Produce Safety and Microbiology Research Unit, Agricultural Research Service, U.S. Department of Agriculture, Albany, California, USA

ABSTRACT *Arcobacter* species are often recovered from marine environments and are isolated from both seawater and shellfish. *Arcobacter marinus* was recovered from the homogenate of a sample containing surface seawater, seaweed, and a starfish. This study describes the whole-genome sequence of the *A. marinus* type strain JCM 15502 (= CL-S1^T = KCCM 90072^T).

Arcobacter species have been isolated from a diverse group of land animals (1, 2). However, other arcobacters are more free living and are often associated with aquatic marine environments. Four *Arcobacter* species were recovered from either seawater (3–5) or hypersaline lagoon water in the Hawaiian Islands (6). A fifth species, *Arcobacter marinus*, was isolated from the homogenate of a sample containing surface seawater from the Sea of Japan (East Sea), seaweed, and a starfish (7). In this study, we report the first closed genome sequence of the *A. marinus* type strain JCM 15502 (= CL-S1^T = KCCM 90072^T).

A. marinus strain JCM 15502^T was grown under the same conditions as *Arcobacter mytili* (8), and genomic DNA was isolated as described (9). Sequencing was first performed on a Roche GS-FLX+ instrument, with libraries constructed using standard protocols. Shotgun and paired-end 454 reads were assembled (Newbler; Roche, version 2.6) into a single chromosomal scaffold of 28 unique contigs. Unscaffolded contigs that were determined to be of low quality (i.e., small contigs containing multiple bases with a quality score of <20 and comprised generally of <20 reads) were deleted. The Perl script contig_extender3 (9) was used to place the remaining 16 contigs at one or more locations within the scaffold. PacBio and Illumina MiSeq sequencing were also performed using standard protocols, with libraries constructed as described (9). The PacBio chromosomal contig and the 454 scaffold contigs were assembled together using SeqMan Pro (version 8.0.2; DNASTAR, Madison, WI), with the remaining 454 contigs added to the assembly manually, using the placement determined above. The Illumina MiSeq reads were assembled using Newbler version 2.6 with default settings. These contigs were also quality controlled (10) and added to the 454/PacBio SeqMan assembly. The MiSeq reads and contigs were used to error correct base calls within the JCM 15502^T sequence, in the same manner as described previously for HiSeq reads (10). Briefly, base calls within contigs at a single location within the assembly were adjusted to the Illumina consensus sequence; single nucleotide polymorphisms within each repeat contig and sequences between the Illumina contigs were addressed by assembling the MiSeq reads onto these regions using Geneious (version 8.1; Biomatters, Auckland, New Zealand) and using the "find variations/SNPs" module, with a default minimum variant frequency parameter of 0.3. The final coverage across the genome was 611×. Chromosomal assembly was also validated using an optical restriction map (restriction enzyme AflII; OpGen, Gaithersburg, MD).

Sequencing metrics and genomic data for *A. marinus* strain JCM 15502^T are presented in Table 1. Strain JCM 15502^T has a circular genome of 2,917,098 bp, with an

Received 14 September 2018 **Accepted** 3 October 2018 **Published** 25 October 2018

Citation Miller WG, Yee E, Huynh S, Parker CT. 2018. Complete genome sequence of the *Arcobacter marinus* type strain JCM 15502. Microbiol Resour Announc 7:e01269-18. https://doi.org/10.1128/MRA.01269-18.

Editor David A. Baltrus, University of Arizona

Address correspondence to William G. Miller, william.miller@ars.usda.gov.

TABLE 1 Sequencing metrics and genomic features of *A. marinus* strain JCM 15502^T

Feature[a]	Value(s)[b]
Sequencing metrics	
454 (shotgun) platform	
No. of reads	255,668
No. of bases	115,346,816
Average length (bases)	451.2
Coverage (×)	39.5
454 (paired-end) platform	
No. of reads	58,215
No. of bases	21,224,999
Average length (bases)	364.6
Coverage (×)	7.3
Illumina MiSeq platform	
No. of reads	2,524,304
No. of bases	746,008,863
Average length (bases)	296
Coverage (×)	255.7
PacBio platform	
No. of reads	89,384
No. of bases	900,973,905
Average length (bases)	10,079.8
Coverage (×)	308.9
Genomic data	
Chromosome	
Size (bp)	2,917,098
G+C content (%)	27.19
No. of CDS[c]	2,717
Assigned function (% CDS)	1,046 (38.5)
General function annotation (% CDS)	1,077 (39.6)
Domain/family annotation only (% CDS)	181 (6.7)
Hypothetical (% CDS)	413 (15.2)
No. of pseudogenes	25
Genomic islands/CRISPR	
No. of genetic islands	3
No. of CDS in genetic islands	78, [1]
No. of CRISPR/Cas loci	0
Gene content/pathways	
Signal transduction	
Che proteins	*cheABCDRVW*(*Y*)$_3$
No. of methyl-accepting chemotaxis proteins	29, [1]
No. of response regulators	59
No. of histidine kinases	75
No. of response regulator/histidine kinase fusions	3
No. of diguanylate cyclases	21
No. of diguanylate phosphodiesterases (HD-GYP, EAL)	7, 6
No. of diguanylate cyclase/phosphodiesterases	12
No. of other	14
Motility	
Flagellin genes	*fla1* to *fla7*
Restriction/modification	
No. of type I systems (*hsd*)	1
No. of type II systems	1
No. of type III systems	0
Transcription/translation	
No. of transcriptional regulatory proteins	73
Non-ECF σ factors	σ^{70}
No. of ECF σ factors	0
No. of tRNAs	63
No. of ribosomal loci	6
CO dehydrogenase (*coxSLF*)	No
Ethanolamine utilization (*eutBCH*)	No
Nitrogen fixation (*nif*)	No
Osmoprotection	BCCT$_5$, *cai/fix*, *betA*, *ectABC*
Pyruvate to acetyl-CoA	
Pyruvate dehydrogenase (E1/E2/E3)	Yes
Pyruvate:ferredoxin oxidoreductase	*por*
Urease	No
Vitamin B$_{12}$ biosynthesis	Yes

[a]CDS, coding sequences; ECF, extracytoplasmic function; acetyl-CoA, acetyl coenzyme A.
[b]Numbers in square brackets indicate pseudogenes or fragments.
[c]Numbers do not include pseudogenes.

average GC content of 27.2%. Protein-, rRNA-, and tRNA-encoding genes were identified (11) and annotated (12) as described. The genome is predicted to encode 2,717 putative protein-coding genes, 25 pseudogenes, 6 rRNA operons, and 63 tRNA-encoding genes. Three genomic islands (19.5, 26.6, and 35.3 kb) were identified in the JCM 15502^T chromosome, with the 35.3-kb island putatively encoding a type VI secretion system.

A noteworthy feature of the *A. marinus* genome is the presence of genes associated with DNA phosphorothioation, in which a nonbridging oxygen in the DNA sugar-phosphate backbone is replaced with sulfur, forming a phosphorothioate linkage (13). In *Streptomyces* spp., this modification involves the *dndABCDE* genes (14). Although *dndBCDE* orthologs are present in strain JCM 15502^T, the cysteine desulfurase *dndA* was not identified; however, in *Escherichia coli*, IscS (present in *A. marinus*) provides the cysteine desulfurase activity for phosphorothioation (15). The function of phosphorothioation in *A. marinus* is unknown but is linked to restriction/modification (16) and resistance to reactive oxidative species (17) in other organisms.

Data availability. The complete genome sequence of *A. marinus* strain JCM 15502^T was deposited in GenBank under the accession number CP032101. The 454, MiSeq, and PacBio sequencing reads were deposited in the NCBI Sequence Read Archive (SRA; accession number SRP155050).

ACKNOWLEDGMENTS

This work was funded by the United States Department of Agriculture, Agricultural Research Service, CRIS projects 2030-42000-230-047 and 2030-42000-230-051.

We thank Maria Figueras for providing *A. marinus* strain JCM 15502^T.

REFERENCES

1. Hsu TT, Lee J. 2015. Global distribution and prevalence of *Arcobacter* in food and water. Zoonoses Public Health 62:579–589. https://doi.org/10.1111/zph.12215.
2. Wesley IV, Miller WG. 2010. *Arcobacter*: an opportunistic human foodborne pathogen? p 185–212. *In* Scheld WM, Grayson ML, Hughes JM (ed), Emerging infections, vol 9. ASM Press, Washington, DC.
3. Levican A, Rubio-Arcos S, Martinez-Murcia A, Collado L, Figueras MJ. 2015. *Arcobacter ebronensis* sp. nov. and *Arcobacter aquimarinus* sp. nov., two new species isolated from marine environment. Syst Appl Microbiol 38:30–35. https://doi.org/10.1016/j.syapm.2014.10.011.
4. Park S, Jung YT, Kim S, Yoon JH. 2016. *Arcobacter acticola* sp. nov., isolated from seawater on the East Sea in South Korea. J Microbiol 54:655–659. https://doi.org/10.1007/s12275-016-6268-4.
5. Zhang Z, Yu C, Wang X, Yu S, Zhang XH. 2016. *Arcobacter pacificus* sp. nov., isolated from seawater of the South Pacific Gyre. Int J Syst Evol Microbiol 66:542–547. https://doi.org/10.1099/ijsem.0.000751.
6. Donachie SP, Bowman JP, On SL, Alam M. 2005. *Arcobacter halophilus* sp. nov., the first obligate halophile in the genus *Arcobacter*. Int J Syst Evol Microbiol 55:1271–1277. https://doi.org/10.1099/ijs.0.63581-0.
7. Kim HM, Hwang CY, Cho BC. 2010. *Arcobacter marinus* sp. nov. Int J Syst Evol Microbiol 60:531–536. https://doi.org/10.1099/ijs.0.007740-0.
8. Miller WG, Yee E, Bono JL. 2018. Complete genome sequence of the Arcobacter mytili type strain LMG 24559. Microbiol Resour Announc 7:e01078-18. https://doi.org/10.1128/MRA.01078-18.
9. Miller WG, Yee E, Lopes BS, Chapman MH, Huynh S, Bono JL, Parker CT, Strachan NJ, Forbes KJ. 2017. Comparative genomic analysis identifies a *Campylobacter* clade deficient in selenium metabolism. Genome Biol Evol 9:1843–1858. https://doi.org/10.1093/gbe/evx093.
10. Miller WG, Yee E. 2018. Complete genome sequence of the Arcobacter trophiarum type strain LMG 25534. Microbiol Resour Announc 7:e01110-18. https://doi.org/10.1128/MRA.01110-18.
11. Miller WG, Yee E, Chapman MH, Smith TP, Bono JL, Huynh S, Parker CT, Vandamme P, Luong K, Korlach J. 2014. Comparative genomics of the *Campylobacter lari* group. Genome Biol Evol 6:3252–3266. https://doi.org/10.1093/gbe/evu249.
12. Miller WG, Yee E, Bono JL. 2018. Complete genome sequence of the Arcobacter bivalviorum type strain LMG 26154. Microbiol Resour Announc 7:e01076-18. https://doi.org/10.1128/MRA.01076-18.
13. Wang L, Chen S, Xu T, Taghizadeh K, Wishnok JS, Zhou X, You D, Deng Z, Dedon PC. 2007. Phosphorothioation of DNA in bacteria by *dnd* genes. Nat Chem Biol 3:709–710. https://doi.org/10.1038/nchembio.2007.39.
14. Zhou X, He X, Liang J, Li A, Xu T, Kieser T, Helmann JD, Deng Z. 2005. A novel DNA modification by sulphur. Mol Microbiol 57:1428–1438. https://doi.org/10.1111/j.1365-2958.2005.04764.x.
15. An X, Xiong W, Yang Y, Li F, Zhou X, Wang Z, Deng Z, Liang J. 2012. A novel target of IscS in *Escherichia coli*: participating in DNA phosphorothioation. PLoS One 7:e51265. https://doi.org/10.1371/journal.pone.0051265.
16. Chen C, Wang L, Chen S, Wu X, Gu M, Chen X, Jiang S, Wang Y, Deng Z, Dedon PC, Chen S. 2017. Convergence of DNA methylation and phosphorothioation epigenetics in bacterial genomes. Proc Natl Acad Sci U S A 114:4501–4506. https://doi.org/10.1073/pnas.1702450114.
17. Wu T, Huang Q, Wang XL, Shi T, Bai L, Liang J, Wang Z, Deng Z, Zhao YL. 2017. Mechanistic investigation on ROS resistance of phosphorothioated DNA. Sci Rep 7:42823. https://doi.org/10.1038/srep42823.