Complete Genome Sequence of the *Arcobacter bivalviorum* Type Strain LMG 26154

William G. Miller,[a] Emma Yee,[a] James L. Bono[b]

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

[b]Meat Safety and Quality Research Unit, U.S. Department of Agriculture, Agricultural Research Service, Clay Center, Nebraska, USA

ABSTRACT   Arcobacters are routinely recovered from marine environments, and multiple *Arcobacter* species have been isolated from shellfish. *Arcobacter bivalviorum* was recovered from mussels collected in the Ebro Delta in northeastern Spain. This report describes the complete whole-genome sequence of the *A. bivalviorum* type strain LMG 26154 (= F4^T = CECT 7835^T).

Members of the genus *Arcobacter* have been isolated from a wide range of hosts and environments (1–3). While many *Arcobacter* species are routinely isolated from terrestrial food animals, at least seven *Arcobacter* species have been recovered from shellfish (4–13). *Arcobacter bivalviorum* is a catalase-positive, urease-negative arcobacter that was isolated originally from mussels collected in the Ebro Delta in Catalonia, Spain (6). The *A. bivalviorum* type strain LMG 26154 (= F4^T = CECT 7835^T) was isolated in March 2007 from mussels of the Ebro Delta. In this study, we report the first closed genome sequence of the *A. bivalviorum* type strain LMG 26154.

Arcobacter bivalviorum was grown at 28°C aerobically for 48 h on anaerobe basal agar (Oxoid) amended with 5% horse blood, and genomic DNA was extracted as described (14) from a 5-μl loop of cells. Shotgun and paired-end Roche 454 reads were generated as described (14) and assembled into 40 contigs using Newbler (version 2.6) with default parameters. A single chromosomal scaffold was produced from 21 unique contigs and closed into a circular, contiguous sequence by using the 19 remaining contigs that represent repeat regions, the Perl script contig_extender3 (14), and directed PCR amplification followed by Sanger sequencing. PacBio sequencing was performed as described (14), generating a single circular chromosomal sequence which was inserted into a SeqMan version 8.0.2 (DNASTAR, Madison, WI) assembly comprised of 454 and Sanger reads, creating a composite 454/PacBio chromosomal sequence and assembly, and additionally verifying the order of the 454 contigs within the original assembly. Base calls within this composite 454/PacBio chromosomal sequence were verified using Illumina HiSeq reads (SeqWright, Houston, TX), as follows. The HiSeq reads were also assembled within Newbler as described above; the resulting trimmed contigs, with a per-base quality score of ≥40, were assembled onto the SeqMan 454/PacBio assembly. Sequences between the Illumina contigs were queried for single nucleotide polymorphisms (SNPs) using the HiSeq reads and the "find variations/SNPs" module within Geneious version 11.0.2 (Biomatters Ltd., Auckland, New Zealand). The final coverage across the closed genome (i.e., a single circular chromosome with no gaps or ambiguous bases) was 1,327×. Chromosomal assembly was also validated using an optical restriction map (restriction enzyme XbaI; OpGen, Gaithersburg, MD).

A summary of the LMG 26154^T genome features is presented in Table 1. *A. bivalviorum* strain LMG 26154^T has a circular genome of 2,684,688 bp with an average GC content of 28.1%. Protein-, rRNA-, and tRNA-encoding genes were identified and

TABLE 1 Sequencing metrics and genomic data for *A. bivalviorum* strain LMG 26154^T

Feature	Value(s)^a
Sequencing metrics	
454 (shotgun) platform	
No. of reads	117,700
No. of bases	65,973,662
Average length (bases)	560.5
Coverage ($\times$)	24.6
454 (paired end) platform	
No. of reads	89,211
No. of bases	28,498,539
Average length (bases)	319.5
Coverage ($\times$)	10.6
Illumina HiSeq platform	
No. of reads	18,785,810
No. of bases	1,878,581,000
Average length (bases)	100
Coverage ($\times$)	699.8
PacBio platform	
No. of reads	356,337
No. of bases	1,588,024,540
Average length (bases)	4,456.5^b
Coverage ($\times$)	591.5
Genomic data	
Chromosome	
Size (bp)	2,684,688
G+C content (%)	28.11
No. of CDS^c	2,584
Assigned function (% CDS)	961 (37.2)
General function annotation (% CDS)	1,018 (39.4)
Domain/family annotation only (% CDS)	180 (7.0)
Hypothetical (% CDS)	425 (16.4)
Pseudogenes	16
Genomic islands/CRISPRs	
No. of genetic islands	4
No. of CDS in genetic islands	49
No. of CRISPR-Cas loci	0
Gene content/pathways	
Signal transduction	
Che proteins	*cheABCDRVW(Y)*$_2$
No. of methyl-accepting chemotaxis proteins	23
No. of response regulators	55
No. of histidine kinases	67, [1]
No. of response regulator/histidine kinase fusions	6
No. of diguanylate cyclases	30, [1]
No. of diguanylate phosphodiesterases (HD-GYP, EAL)	9, 5
No. of diguanylate cyclase/phosphodiesterases	12
No. of other	10
Motility	
Flagellin genes	*flaAB*
Restriction/modification	
No. of type I systems (*hsd*)	0
No. of type II systems	1
No. of type III systems	0
Transcription/translation	
No. of transcriptional regulatory proteins	64, [1]
Non-ECF^d σ factor	σ^{70}
No. of ECF σ factors	2
No. of tRNAs	47
No. of ribosomal loci	4
Nitrogen fixation (*nif*)	No
Osmoprotection	*ectABC, proVWX*
Pyruvate $\rightarrow$ acetyl-CoA	
Pyruvate dehydrogenase (E1/E2/E3)	Yes
Pyruvate:ferredoxin oxidoreductase	*por*
Urease	No
Vitamin B$_{12}$ biosynthesis	Yes

^{a}Numbers in brackets indicate pseudogenes or fragments.
^{b}Maximum length, 29,273 bp.
^{c}Numbers do not include pseudogenes. CDS, coding sequences.
^{d}ECF, extracytoplasmic function.

annotated as described (15), with the exception that the composite proteome used here for comparison included proteins derived from all currently completed *Campylobacter* and *Arcobacter* genomes. The genome is predicted to encode 2,584 putative protein-coding genes and 16 pseudogenes. Additionally, the LMG 26154^T genome contains 4 rRNA operons and 47 tRNA-encoding genes. Four genomic islands were identified in the LMG 26154^T chromosome. Two of these genomic islands are approximately 10 kb in size, and two are approximately 15 kb in size. No plasmids were identified in strain LMG 26154^T.

Strain LMG 26154^T contains 25 of the 30 genes necessary for *de novo* biosynthesis of adenosylcobalamin (coenzyme B$_{12}$) via the oxygen-independent (anaerobic) pathway (16). Orthologs of *cbiJ*, *cobA*, *cobC*, *fre*, and *pduX* were not identified; however, these enzymatic functions could be provided by other proteins or by orthologs with low similarity to proteins currently in the NCBI nr database.

Data availability. The complete genome sequence of *A. bivalviorum* strain LMG 26154^T has been deposited in GenBank under the accession number CP031217. Illumina HiSeq, 454, and PacBio sequencing reads have been deposited in the NCBI Sequence Read Archive (SRA) under the accession number SRP154993.

REFERENCES

1. Collado L, Figueras MJ. 2011. Taxonomy, epidemiology, and clinical relevance of the genus *Arcobacter*. Clin Microbiol Rev 24:174–192. https://doi.org/10.1128/CMR.00034-10.
2. Forsythe SJ, 2006. *Arcobacter*, p. 181–221. *In* Motarjemi Y, Adams M (ed), Emerging foodborne pathogens. Woodhead Publishing Ltd, Cambridge, England.
3. Wesley IV, Miller WG, 2010. *Arcobacter*: an opportunistic human foodborne pathogen? p. 185–212. *In* Scheld WM, Grayson ML, Hughes JM (ed), Emerging Infections 9. ASM Press, Washington, DC.
4. Mizutani Y, Tanaka R. 2017. Genome sequence of *Arcobacter* sp. strain LA11, isolated from the abalone *Haliotis discus*. Genome Announc 5:e00032-17. https://doi.org/10.1128/genomeA.00032-17.
5. Tanaka R, Cleenwerck I, Mizutani Y, Iehata S, Bossier P, Vandamme P. 2017. *Arcobacter haliotis* sp. nov., isolated from abalone species *Haliotis gigantea*. Int J Syst Evol Microbiol 67:3050–3056. https://doi.org/10.1099/ijsem.0.002080.
6. Levican A, Collado L, Aguilar C, Yustes C, Diéguez AL, Romalde JL, Figueras MJ. 2012. *Arcobacter bivalviorum* sp. nov. and *Arcobacter venerupis* sp. nov., new species isolated from shellfish. Syst Appl Microbiol 35:133–138. https://doi.org/10.1016/j.syapm.2012.01.002.
7. Morejon IF, González A, Ferrús MA. 2017. Detection, identification, and antimicrobial susceptibility of *Arcobacter* spp. isolated from shellfish in Spain. Foodborne Pathog Dis 14:238–243. https://doi.org/10.1089/fpd.2016.2202.
8. Diéguez AL, Balboa S, Magnesen T, Romalde JL. 2017. *Arcobacter lekithochrous* sp. nov., isolated from a molluscan hatchery. Int J Syst Evol Microbiol 67:1327–1332. https://doi.org/10.1099/ijsem.0.001809.
9. Collado L, Cleenwerck I, Van Trappen S, De Vos P, Figueras MJ. 2009. *Arcobacter mytili* sp. nov., an indoxyl acetate-hydrolysis-negative bacterium isolated from mussels. Int J Syst Evol Microbiol 59:1391–1396. https://doi.org/10.1099/ijs.0.003749-0.
10. Figueras MJ, Collado L, Levican A, Perez J, Solsona MJ, Yustes C. 2011. *Arcobacter molluscorum* sp. nov., a new species isolated from shellfish. Syst Appl Microbiol 34:105–109. https://doi.org/10.1016/j.syapm.2010.10.001.
11. Figueras MJ, Levican A, Collado L, Inza MI, Yustes C. 2011. *Arcobacter ellisii* sp. nov., isolated from mussels. Syst Appl Microbiol 34:414–418. https://doi.org/10.1016/j.syapm.2011.04.004.
12. Bonerba E, Mottola A, Parisi A, Di Pinto A, Serraino A, Bozzo G, Giacometti F, Ceci E, Tantillo G. 2015. Detection of *Arcobacter* spp. in *Mytilus galloprovincialis* samples collected from Apulia region. Ital J Food Saf 4:4583. https://doi.org/10.4081/ijfs.2015.4583.
13. 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.
14. 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.
15. 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.
16. Fang H, Kang J, Zhang D. 2017. Microbial production of vitamin B$_{12}$: a review and future perspectives. Microb Cell Fact 16:15. https://doi.org/10.1186/s12934-017-0631-y.