AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of the *Arcobacter ellisii* Type Strain LMG 26155

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

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

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

ABSTRACT *Arcobacter* spp. are highly prevalent in contaminated environmental waters and have been recovered from both freshwater and seawater, with several species isolated from shellfish. *Arcobacter ellisii* was recovered from mussels collected in Catalonia, Spain. This study describes the whole-genome sequence of the *A. ellisii* type strain LMG 26155 (=F79-6^T =CECT 7837^T).

Arcobacter species are a common contaminant of freshwater and seawater (1–3) and molluscs (4–8). *Arcobacter ellisii* was isolated during a survey in Catalonia, Spain, to detect *Arcobacter* spp. in shellfish (9). Although isolated originally from mussels, *A. ellisii* has also been detected at a spinach processing plant (10) and at a wastewater treatment plant (11), suggesting that this species might be a general water contaminant that concentrates in mussels through filter feeding. In this study, we report the first closed genome sequence of the *A. ellisii* type strain LMG 26155 (=F79-6^T =CECT 7837^T), isolated from mussels of the Ebro Delta.

Arcobacter ellisii strain LMG 26155^T was grown aerobically at 30°C for 48 h on anaerobe basal agar (Oxoid) amended with 5% horse blood. A loop (~5 μl) of cells was harvested, and genomic DNA was isolated using the Promega Wizard kit, as described previously (12). Shotgun and paired-end Roche 454 reads were generated on a GS-FLX+ instrument, with the Titanium chemistry and standard protocols; these were assembled into 77 contigs and a single chromosomal scaffold containing 27 unique contigs using Newbler (version 2.6; Roche), with default settings. Contigs of low quality that were composed of <100 reads were deleted. The Perl script contig_extender3 (12) was used to place within the chromosomal scaffold the 25 remaining contigs that mostly represent repeat regions. Library preparation and sequencing on the PacBio RS II platform (Pacific Biosciences, Menlo Park, CA) were performed as described previously (12), using standard methodology. RS_HGAP_Assembly v3 (Pacific Biosciences) with default settings was used to assemble the reads. A single chromosomal contig was obtained which was subjected to one round of polishing using RS_resequencing.1 and quality trimmed and circularized within Geneious (version 8.0; Biomatters Ltd., Auckland, New Zealand). The PacBio contig was assembled along with the 454 contigs using SeqMan Pro (version 8.0.2; DNAStar, Madison, WI), with the unique 454 contigs assembled onto the PacBio contig automatically and the repeat 454 contigs added in manually. Base calls within the composite 454/PacBio assembly were verified using Illumina HiSeq reads (SeqWright, Houston, TX), as described previously (13). The final coverage across the genome was 925×. Chromosomal assembly was also validated using an optical restriction map (restriction enzyme XbaI; OpGen, Gaithersburg, MD).

The LMG 26155^T genome features are summarized in Table 1. *A. ellisii* strain LMG 26155^T has a circular genome of 2,799,949 bp, with an average GC content of 26.9%. Protein-, rRNA-, and tRNA-encoding genes were identified and annotated as described

Received 14 September 2018 **Accepted** 30 September 2018 **Published** 25 October 2018

Citation Miller WG, Yee E, Bono JL. 2018. Complete genome sequence of the *Arcobacter ellisii* type strain LMG 26155. Microbiol Resour Announc 7:e01268-18. https://doi.org/10.1128/MRA.01268-18.

Editor David A. Baltrus, University of Arizona

This is a work of the U.S. Government and is not subject to copyright protection in the United States. Foreign copyrights may apply.

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

TABLE 1 Sequencing metrics and genomic data for *A. ellisii* strain LMG 26155^T

Feature[a]	Value(s)[b]
Sequencing metrics	
454 (shotgun) platform	
No. of reads	141,497
No. of bases	87,366,368
Avg length (bases)	617.4
Coverage (×)	31.2
454 (paired-end) platform	
No. of reads	78,872
No. of bases	24,957,376
Avg length (bases)	316.4
Coverage (×)	8.9
Illumina HiSeq 2000 platform	
No. of reads	15,212,786
No. of bases	1,529,278,600
Avg length (bases)	101
Coverage (×)	546.2
PacBio platform	
No. of reads	328,963
No. of bases	947,635,645
Avg length (bases)	2,880.7[c]
Coverage (×)	338.4
Genomic data	
Chromosome	
Size (bp)	2,799,949
G+C content (%)	26.92
No. (%) of CDS[d]	2,743
Assigned function	973 (35.5)
General function annotation	1,049 (38.2)
Domain/family annotation only	178 (6.5)
Hypothetical	543 (19.8)
Pseudogenes	32
Genomic islands/CRISPR	
No. of genetic islands	8
No. of CDS in genetic islands	157, [2]
CRISPR/Cas loci	III-B
Gene content/pathways	
IS elements, mobile elements, or transposases	6, [1] (IS*3*); 2 (IS*4*); 4 (IS*21*); 1 (IS*1249*); 1 (IS*1634*); 1 (other)
Signal transduction	
Che proteins	*cheABCDRVW*$(Y)_2$
No. of methyl-accepting chemotaxis proteins	26
No. of response regulators	51, [1]
No. of histidine kinases	60, [2]
No. of response regulator/histidine kinase fusions	7
No. of diguanylate cyclases	24
No. of diguanylate phosphodiesterases (HD-GYP, EAL)	8, 4
No. of diguanylate cyclase/phosphodiesterases	14
No. of other proteins	13
Motility	
Flagellin gene	*fla*
Restriction/modification (no.)	
Type I systems (*hsd*)	3
Type II systems	0
Type III systems	0
Transcription/translation	
No. of transcriptional regulatory proteins	60, [1]
Non-ECF σ factors	σ^{54}, σ^{70}
No. of ECF σ factors	0
No. of tRNAs	60
No. of ribosomal loci	5
CO dehydrogenase (*coxSLF*)	Yes
Ethanolamine utilization (*eutBCH*)	No
Nitrogen fixation (*nif*)	Yes
Osmoprotection	BCCT, *betA*, *ectABCD*
Pyruvate → acetyl-CoA	
Pyruvate dehydrogenase (E1/E2/E3)	No
Pyruvate:ferredoxin oxidoreductase	*porABDG*
Urease	*ureAB*, *ucaA*
Vitamin B_{12} biosynthesis	No

[a]CDS, coding sequences; ECF, extracytoplasmic function; CoA, coenzyme A.
[b]Numbers in square brackets indicate pseudogenes or fragments. BCCT, betaine-carnitine-choline transporter.
[c]Maximum length, 25,296 bases.
[d]Numbers do not include pseudogenes.

previously (14, 15). The genome is predicted to contain 2,743 putative protein-coding genes, 32 pseudogenes, five rRNA operons, and 60 tRNA-encoding genes. A type III-B CRISPR/Cas system and at least eight genomic islands, ranging in size from 5.8 kb to 37.5 kb, were identified in the LMG 26155^T genome. Fourteen insertion sequences (IS) from the following families are present in the *A. ellisii* chromosome: IS*3* ($n = 6$), IS*4* ($n = 2$), IS*21* ($n = 4$), IS*1249* ($n = 1$), and IS*1634* ($n = 1$). Similar to *Arcobacter mytili* (16), a large type 1 secretion system (T1SS) repeat domain-containing protein-encoding gene with a tandem-repeat internal motif (29 × 414 bp) was identified.

Data availability. The complete genome sequence of *A. ellisii* strain LMG 26155^T has been deposited in GenBank under the accession number CP032097. The 454, HiSeq, and PacBio sequencing reads have been deposited in the NCBI Sequence Read Archive (SRA; accession number SRP155046).

ACKNOWLEDGMENTS

This work was funded by the U.S. Department of Agriculture, Agricultural Research Service, CRIS projects 2030-42000-230-047, 2030-42000-230-051, and 3040-42000-015-00D.

We thank Maria Figueras for providing *A. ellisii* strain LMG 26155^T.

REFERENCES

1. Forsythe SJ. 2006. *Arcobacter*, p 181–221. *In* Motarjemi Y, Adams M (ed), Emerging foodborne pathogens. Woodhead Publishing Ltd., Cambridge, England.
2. 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.
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. 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.
5. 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.
6. Levican A, Collado L, Aguilar C, Yustes C, Dieguez 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. 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.
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. 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.
10. Hausdorf L, Neumann M, Bergmann I, Sobiella K, Mundt K, Frohling A, Schluter O, Klocke M. 2013. Occurrence and genetic diversity of *Arcobacter* spp. in a spinach-processing plant and evaluation of two *Arcobacter*-specific quantitative PCR assays. Syst Appl Microbiol 36:235–243. https://doi.org/10.1016/j.syapm.2013.02.003.
11. Levican A, Collado L, Figueras MJ. 2016. The use of two culturing methods in parallel reveals a high prevalence and diversity of *Arcobacter* spp. in a wastewater treatment plant. Biomed Res Int 2016:8132058. https://doi.org/10.1155/2016/8132058.
12. 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.
13. 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.
14. 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.
15. 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.
16. 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.