



Complete Genome Sequence of the *Arcobacter halophilus* Type Strain CCUG 53805

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

^aProduce Safety and Microbiology Research Unit, Agricultural Research Service, U.S. Department of Agriculture, Albany, California, USA

^bMeat Safety and Quality Research Unit, Agricultural Research Service, U.S. Department of Agriculture, Clay Center, Nebraska, USA

ABSTRACT Many *Arcobacter* spp. are free living and are routinely recovered from marine environments. *Arcobacter halophilus* was isolated from hypersaline lagoon water in the Hawaiian islands, and it was demonstrated to be an obligate halophile. This study describes the complete whole-genome sequence of the *A. halophilus* type strain, CCUG 53805 (= LA31B^T = ATCC BAA-1022^T).

A *Arcobacter* species are often recovered from marine environments. Although many *Arcobacter* taxa are isolated from shellfish (1–4), others are free living and have been recovered from seawater (5, 6) or marine sediments (7–9). *Arcobacter halophilus* is an obligate halophile that was cultivated from a water sample collected in October 2000 from the hypersaline Lake Laysan at Laysan Atoll in the Northwestern Hawaiian Islands (10, 11). In this study, we report the first closed genome sequence of the *A. halophilus* type strain, CCUG 53805 (= LA31B^T = ATCC BAA-1022^T).

Arcobacter halophilus CCUG 53805^T was grown aerobically for 48 h at 30°C on anaerobe basal agar (Oxoid) amended with 5% horse blood and 2% (wt/vol) NaCl. Genomic DNA was prepared from a loop of cells as described previously (12). Shotgun and paired-end Roche GS-FLX+ reads were assembled using Newbler v. 2.6, yielding 91 total contigs and a chromosomal scaffold of 45 unique contigs. Forty-six contigs, representing regions present more than once in the chromosome, were positioned into the scaffold gaps using the custom Perl script *contig_extender3* (12). Contig junctions and any remaining small gaps were validated and closed using directed PCR amplification/Sanger sequencing. Joining of the 454 contigs and linking Sanger sequences into a single chromosomal sequence, using the contig order obtained above, was performed using SeqMan v. 8.0.2 (DNASTAR, Madison, WI). The contig order within the 454 sequence was also verified using an optical restriction map (restriction enzyme *SpeI*; OpGen, Gaithersburg, MD). During closure, a large repetitive region within the chromosome was identified, and PacBio reads that spanned this region were generated. PacBio sequencing was performed as described (12) and generated a single circular sequence that was added to the 454 SeqMan assembly, further confirming the contig order within the 454 scaffold. Illumina HiSeq reads (SeqWright, Houston, TX) independently verified all base calls within the chromosome. These reads were assembled *de novo* using Newbler, and the resulting contigs were assembled onto the SeqMan 454/PacBio sequence as described (12), with the PacBio sequence corrected with respect to the HiSeq consensus; single-nucleotide polymorphisms (SNPs) in the sequences between the HiSeq contigs were identified using the Geneious v. 11.0.2 (Biomatters Ltd., Auckland, New Zealand) “find variations/SNPs” module with the default minimum variant frequency parameter of 0.3. The final coverage across the genome was 919×.

Genome feature data for *A. halophilus* strain CCUG 53805^T are presented in Table 1. The CCUG 53805^T genome is 2,812,536 bp, with an average G+C content of 27.6%. Protein-, rRNA-, and tRNA-encoding genes were identified and annotated as described

Received 2 August 2018 Accepted 12 September 2018 Published 11 October 2018

Citation Miller WG, Yee E, Bono JL. 2018. Complete genome sequence of the *Arcobacter halophilus* type strain CCUG 53805. Microbiol Resour Announc 7:e01077-18. <https://doi.org/10.1128/MRA.01077-18>.

Editor Iddo Friedberg, Iowa State University

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. halophilus* strain CCUG 53805^T

Feature	Data ^a
Sequencing metric	
Platform	
454 (shotgun)	
No. of reads	158,389
No. of bases	53,780,918
Average length (bases)	340
Coverage (×)	19.1
454 (paired end)	
No. of reads	670,813
No. of bases	215,687,983
Average length (bases)	322
Coverage (×)	95.8
Illumina HiSeq 2000	
No. of reads	18,199,888
No. of bases	1,838,188,068
Average length (bases)	101
Coverage (×)	653.6
PacBio	
No. of reads	125,464
No. of bases	476,556,540
Average length (bases)	3,798.4 ^b
Coverage (×)	169.4
Genomic data	
Chromosome	
Size (bp)	2,812,536
G+C content (%)	27.61
No. of CDS ^c	2,622
Assigned function (% CDS)	1,009 (38.5)
General function annotation (% CDS)	1,025 (39.1)
Domain/family annotation only (% CDS)	173 (6.6)
Hypothetical (% CDS)	415 (15.8)
No. of pseudogenes	19
Genomic islands/CRISPR	
No. of genetic islands	6
No. of CDS in genetic islands	106 [1]
CRISPR/Cas loci	I-B
Gene content/pathways	
Signal transduction	
Che proteins	<i>cheABCDRVW</i> (Y) ₃
No. of methyl-accepting chemotaxis proteins	29
No. of response regulators	60
No. of histidine kinases	76
No. of response regulator/histidine kinase fusions	3
No. of diguanylate cyclases	26
No. of diguanylate phosphodiesterases (HD-GYP, EAL)	5, 5
No. of diguanylate cyclase/phosphodiesterases	11 [1]
No. of others	13
Motility	
Flagellin genes	<i>fla1, fla2, fla3, fla4</i>
Restriction/modification	
No. of type I (<i>hsd</i>) systems	2
No. of type II systems	1
No. of type III systems	0
Transcription/translation	
No. of transcriptional regulatory proteins	60
Non-ECF σ factors ^d	σ^{70}
No. of ECF σ factors	1
No. of tRNAs	63
No. of ribosomal loci	6
Nitrogen fixation (<i>nif</i>)	No
Osmoprotection	BCCT ₄ , <i>cai/fix, betA, ectABCD, proABCWX</i>
Pyruvate → acetyl-CoA ^e	
Pyruvate dehydrogenase (E1/E2/E3)	Yes
Pyruvate:ferredoxin oxidoreductase	<i>por</i>
Urease	No
Vitamin B ₁₂ biosynthesis	No

^aNumbers in brackets indicate pseudogenes/fragments.

^bMaximum length, 23,873 bp.

^cNumbers do not include pseudogenes. CDS, coding DNA sequences.

^dECF, extracytoplasmic function.

^eCoA, coenzyme A.

(13) using the custom Perl script BlastPTrimmer14 (12), with the composite proteome used here for BLAST analyses and comparisons containing proteins from all current complete *Campylobacter* and *Arcobacter* genomes. The genome is predicted to encode 2,622 putative protein-coding genes and 19 pseudogenes. Additionally, the CCUG 53805^T genome contains 6 rRNA operons and 63 tRNA-encoding genes. A type I-B CRISPR/Cas system and 6 genomic islands ranging from 8.5 kb to 26.9 kb were identified in the CCUG 53805^T chromosome; the largest genomic island putatively encodes a type VI secretion system. The CCUG 53805^T genome contains no plasmids.

Consistent with its description as an obligate halophile, the genome of strain CCUG 53805^T contains several genes associated with the biosynthesis and uptake of osmolytes. These include genes for the biosynthesis of ectoine (*ectABC*), 5-hydroxyectoine (*ectD*), proline (*proABC*), glycine betaine aldehyde (*betA*), and γ -butyrobetaine (*cai/fix*). Additionally, the CCUG 53805^T genome encodes the ProVWX proline/glycine betaine ABC transporter and four BCCT (betaine/carnitine/choline transporter) family proteins.

Data availability. The complete genome sequence of *A. halophilus* strain CCUG 53805^T has been deposited in GenBank under the accession number [CP031218](#). The 454, HiSeq, and PacBio sequencing reads have been deposited in the NCBI Sequence Read Archive (SRA) under the accession number [SRP155008](#). The source codes for contig_extender3 and BlastPTrimmer14 are available through GitHub (https://github.com/Arcobacter/Genome_perl).

ACKNOWLEDGMENTS

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

We thank Stephen On for providing *A. halophilus* strain CCUG 53805^T.

REFERENCES

- Dieguez 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>.
- 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>.
- 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>.
- 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>.
- Fera MT, Mauger TL, Gugliandolo C, Beninati C, Giannone M, La Camera E, Carbone M. 2004. Detection of *Arcobacter* spp. in the coastal environment of the Mediterranean Sea. *Appl Environ Microbiol* 70:1271–1276. <https://doi.org/10.1128/AEM.70.3.1271-1276.2004>.
- 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>.
- Llobet-Brossa E, Rosselló-Mora R, Amann R. 1998. Microbial community composition of Wadden Sea sediments as revealed by fluorescence *in situ* hybridization. *Appl Environ Microbiol* 64:2691–2696.
- Bowman JP, McCuaig RD. 2003. Biodiversity, community structural shifts, and biogeography of prokaryotes within Antarctic continental shelf sediment. *Appl Environ Microbiol* 69:2463–2483. <https://doi.org/10.1128/AEM.69.5.2463-2483.2003>.
- Sasi Jyothsna TS, Rahul K, Ramaprasad EVV, Sasikala C, Ramana CV. 2013. *Arcobacter anaerophilus* sp. nov., isolated from an estuarine sediment and emended description of the genus *Arcobacter*. *Int J Syst Evol Microbiol* 63:4619–4625. <https://doi.org/10.1099/ijms.0.054155-0>.
- 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/ijms.0.63581-0>.
- Donachie SP, Hou S, Lee KS, Riley CW, Pikina A, Belisle C, Kempe S, Gregory TS, Bossuyt A, Boerema J, Liu J, Freitas TA, Malahoff A, Alam M. 2004. The Hawaiian Archipelago: a microbial diversity hotspot. *Microb Ecol* 48:509–520. <https://doi.org/10.1007/s00248-004-0217-1>.
- 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>.
- 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>.