AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequences of the *Arcobacter cryaerophilus* Strains ATCC 43158^T and ATCC 49615

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 cryaerophilus* was originally recovered from aborted bovine and porcine fetuses, but it has been subsequently isolated from meat, water, and human clinical samples. This study describes the complete whole-genome sequences of two *A. cryaerophilus* strains, ATCC 43158^T (=A 169/B^T =LMG 24291^T) and ATCC 49615 (=CDC D2610 =LMG 10829).

Arcobacter cryaerophilus (formerly *Campylobacter cryaerophila* [1]) was originally recovered from aborted bovine and porcine fetuses (1–3), from the milk of cows with mastitis (4), and from bovine and porcine feces (1). Subsequent studies reported the recovery of *A. cryaerophilus* isolates from meat (5, 6), water (7), and human clinical samples (8, 9). Additionally, based on the results of whole-cell protein and rRNA restriction fragment length polymorphism analyses (10, 11), *A. cryaerophilus* strains have been divided into two subgroups, 1A and 1B. In this study, we report the first closed genome sequences of two *A. cryaerophilus* strains, the type strain ATCC 43158 (subgroup 1A; =A 169/B^T [1] =LMG 24291^T), isolated in the United Kingdom from the brain of an aborted bovine fetus, and strain ATCC 49615 (subgroup 1B; =CDC D2610 [12] =LMG 10829), isolated in Illinois from human blood. A draft genome sequence (91 scaffolds) was deposited previously for *A. cryaerophilus* strain ATCC 43158^T (GenBank accession number NXGK01000000). Additionally, draft genome sequences for 12 other *A. cryaerophilus* strains have been deposited previously in GenBank (https://www.ncbi.nlm.nih.gov/genome/genomes/11530).

Both *A. cryaerophilus* strains were grown aerobically at 30°C for 48 h on anaerobe basal agar (Oxoid) amended with 5% horse blood. Cells were separately removed from each plate using sterile 5-μl inoculating loops, and genomic DNA was prepared from each loopful of cells using the Promega Wizard genomic DNA purification kit (Madison, WI). Roche 454 shotgun and paired-end libraries were constructed for both strains, using the manufacturer's standard protocols, and sequenced on a GS-FLX+ instrument using the Titanium chemistry. For each strain, the shotgun and paired-end reads were assembled together, using Newbler v. 2.6 (Roche) and default parameters; 454 reads were quality controlled within Newbler. Sequencing metrics for both strains are presented in Table 1. Two scaffolds were obtained for strain ATCC 43158^T, one chromosomal scaffold of 20 contigs and a megaplasmid scaffold of 4 contigs. One chromosomal scaffold of 16 contigs was obtained for strain ATCC 49615. Each scaffold was closed by using the custom Perl script contig_extender3 (13) to place within the scaffold gaps unique, nonscaffolded contigs and/or contigs representing sequences present at two or more locations. The two chromosomes and the megaplasmid were assembled manually within SeqMan Pro v. 8.0 (DNASTAR, Madison, WI), using the contig order defined above and contig-spanning 454 reads. Assembly of both genomes was verified using an optical restriction map (restriction enzyme XbaI; OpGen, Gaith-

Received 23 October 2018 **Accepted** 1 November 2018 **Published** 21 November 2018

Citation Miller WG, Yee E, Bono JL. 2018. Complete genome sequences of the *Arcobacter cryaerophilus* strains ATCC 43158^T and ATCC 49615. Microbiol Resour Announc 7:e01463-18. https://doi.org/10.1128/MRA.01463-18.

Editor Frank J. Stewart, Georgia Institute of Technology

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

TABLE 1 Sequencing metrics and genomic data for *Arcobacter cryaerophilus* strains ATCC 43158^T and ATCC 49615

Feature[b]	Value(s)[a]	
	ATCC 43158^T	ATCC 49615
Sequencing metrics		
454 (shotgun) platform		
No. of reads	215,129	87,759
No. of bases	90,272,116	36,134,765
Avg length (bases)	419.6	411.7
Coverage (×)	45.0	17.6
454 (paired-end) platform		
No. of reads	118,785	161,571
No. of bases	38,763,675	49,505,772
Avg length (bases)	326.3	306.4
Coverage (×)	19.3	24.1
Illumina HiSeq 2000 platform		
No. of reads	18,366,932	18,769,966
No. of bases	1,836,693,200	1,895,766,566
Avg length (bases)	100	101
Coverage (×)	915.2	922.1
PacBio platform		
No. of reads	117,659	NA
No. of bases	437,690,629	NA
Avg length (bases)	3,720[c]	NA
Coverage (×)	218.1	NA
Newbler metrics[d]		
Total no. of contigs	45	32
N50ContigSize (454)	195,432	346,474
Q40PlusBases (454) (%)	99.90	99.87
N50ContigSize (HiSeq pool 1)	195,198	301,090
Q40PlusBases (HiSeq pool 1) (%)	99.98	99.96
N50ContigSize (HiSeq pool 2)	195,524	165,874
Q40PlusBases (HiSeq pool 2) (%)	99.96	99.96
N50ContigSize (HiSeq pool 3)	195,196	301,081
Q40PlusBases (HiSeq pool 3) (%)	99.97	99.99
Genomic data		
Size (bp)	2,006,909; 101,431	2,055,914
G+C content (%)	27.46; 24.63	27.51
No. of CDS[e]	1,951; 94	1,994
Assigned function (% CDS)	847 (43.4); 3 (3.2)	857 (43.0)
General function annotation (% CDS)	675 (34.6); 34 (36.2)	694 (34.8)
Domain/family annotation only (% CDS)	131 (6.7); 4 (4.3)	136 (6.8)
Hypothetical (% CDS)	298 (15.3); 53 (56.4)	307 (15.4)
Pseudogenes	23; 6	8
Genomic islands/CRISPR		
No. of genetic islands	8; 0	5
No. of CDS in genetic islands	108 [2]; 0	85 [2]
No. of CRISPR/Cas loci	0	0
Gene content/pathways		
IS elements, mobile elements, or transposases	2; 1	2
Signal transduction		
No. of Che proteins	9; 0	9
No. of methyl-accepting chemotaxis proteins	20; 0	23
No. of response regulators	23; 2	25
No. of histidine kinases	25; 1	27
No. of response regulator/histidine kinase fusions	0	3
No. of diguanylate cyclases	14; 0	15
No. of diguanylate phosphodiesterases (HD-GYP, EAL)	4, 1; 0	5, 1
No. of diguanylate cyclase/phosphodiesterases	9; 0	9
No. of other	10; 1	10
Motility		
Flagellin genes	*flaAB*	*flaAB*
Restriction/modification		
No. of type I systems (*hsd*)	[2]; 0	2
No. of type II systems	1; 0	3
No. of type III systems	0	2

(Continued on next page)

TABLE 1 (Continued)

Feature[b]	Value(s)[a]	
	ATCC 43158^T	ATCC 49615
Transcription/translation		
No. of transcriptional regulatory proteins	32; 4	32
Non-ECF σ factors	σ^{70}	σ^{70}
No. of ECF σ factors	0	0
No. of tRNAs	49; 0	51
No. of ribosomal loci	5; 0	5
CO dehydrogenase (*coxSLF*)	No	No
Ethanolamine utilization (*eutBCH*)	No	No
Nitrogen fixation (*nif*)	No	No
Osmoprotection	*ectABC*	No
Pyruvate → acetyl-CoA		
Pyruvate dehydrogenase (E1/E2/E3)	Yes	Yes
Pyruvate:ferredoxin oxidoreductase	No	No
Urease	No	No
Vitamin B_{12} biosynthesis	No	No

[a] Numbers in square brackets indicate pseudogenes or fragments. Strain ATCC 43158^T values before a semicolon are for the chromosome, while values after the semicolon are for the pACRY43158 plasmid. NA, not applicable.
[b] CDS, coding sequences; ECF, extracytoplasmic function; acetyl-CoA, acetyl coenzyme A.
[c] Maximum length, 25,867 bases.
[d] Features and values taken from largeContigMetrics within 454NewblerMetrics.txt for each assembly. Large contigs were defined as having ≥500 bases. Due to the large number of HiSeq reads, the total reads were split into three pools and assembled independently. N50ContigSize value is in number of bases.
[e] Numbers do not include pseudogenes.

ersburg, MD). Illumina HiSeq reads were obtained from SeqWright (Houston, TX) and assembled and quality controlled within Newbler v. 2.6, as above. The Illumina contigs were assembled automatically (unique contigs) or positioned manually (repeated contigs) within the 454 SeqMan assembly. The Illumina contigs and reads were used to verify the 454 base calls, as described previously (14). A repeat region within the type strain chromosome required long reads for validation. Therefore, a 20-kb PacBio library was constructed, using standard methods (15), and sequenced on an RS II instrument, using standard protocols, as described previously (14). PacBio reads were assembled with RS_HGAP_Assembly v. 3 (Pacific Biosciences) and default settings, which yielded single chromosomal and megaplasmid contigs. The PacBio chromosomal contig was manually placed within the SeqMan assembly to confirm the 454 contig order across the repeat region.

Genomic data for the two *A. cryaerophilus* strains are presented in Table 1. The average genome size for the two strains is approximately 2 Mbp, with an average G+C content of 27.5%. Protein-, rRNA-, and tRNA-encoding genes were identified using GeneMark, RNAmmer, and ARAGORN (16–18), respectively, and annotated as described previously (14). The ATCC 43158^T genome also contains the 101,431-bp megaplasmid pACRY43158.

Data availability. The complete genome sequence of *A. cryaerophilus* strain ATCC 43158^T has been deposited in GenBank under the accession numbers CP032823 (chromosome) and CP032824 (pACRY43158), and the complete genome sequence of *A. cryaerophilus* strain ATCC 49615 has been deposited in GenBank under the accession number CP032825. All 454, HiSeq, and PacBio sequencing reads for strain ATCC 43158^T and 454 and HiSeq reads for strain ATCC 49615 have been deposited in the NCBI Sequence Read Archive (SRA) under the accession numbers SRP164716 and SRP164722, respectively.

ACKNOWLEDGMENT

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.

REFERENCES

1. Neill SD, Campbell JN, O'Brien JJ, Weatherup STC, Ellis WA. 1985. Taxonomic position of *Campylobacter cryaerophila* sp. nov. Int J Syst Bacteriol 35:342–356. https://doi.org/10.1099/00207713-35-3-342.
2. Higgins R, Degre R. 1979. Isolation of *Spirillum*-like organisms from pig and bovine fetuses. Vet Rec 104:262–263. https://doi.org/10.1136/vr.104.12.262-a.
3. Higgins R, Degre R. 1979. *Spirillum*-like organism isolated from pig and bovine fetuses. Vet Rec 104:559. https://doi.org/10.1136/vr.104.24.559-b.
4. Logan EF, Neill SD, Mackie DP. 1982. Mastitis in dairy cows associated with an aerotolerant campylobacter. Vet Rec 110:229–230. https://doi.org/10.1136/vr.110.10.229.
5. Van Driessche E, Houf K. 2007. Characterization of the *Arcobacter* contamination on Belgian pork carcasses and raw retail pork. Int J Food Microbiol 118:20–26. https://doi.org/10.1016/j.ijfoodmicro.2007.05.006.
6. Zacharow I, Bystroń J, Wałecka-Zacharska E, Podkowik M, Bania J. 2015. Genetic diversity and incidence of virulence-associated genes of *Arcobacter butzleri* and *Arcobacter cryaerophilus* isolates from pork, beef, and chicken meat in Poland. Biomed Res Int 2015:956507. https://doi.org/10.1155/2015/956507.
7. Collado L, Inza I, Guarro J, Figueras MJ. 2008. Presence of *Arcobacter* spp. in environmental waters correlates with high levels of fecal pollution. Environ Microbiol 10:1635–1640. https://doi.org/10.1111/j.1462-2920.2007.01555.x.
8. Mandisodza O, Burrows E, Nulsen M. 2012. *Arcobacter* species in diarrhoeal faeces from humans in New Zealand. N Z Med J 125:40–46.
9. Van den Abeele AM, Vogelaers D, Vanlaere E, Houf K. 2016. Antimicrobial susceptibility testing of *Arcobacter butzleri* and *Arcobacter cryaerophilus* strains isolated from Belgian patients. J Antimicrob Chemother 71:1241–1244. https://doi.org/10.1093/jac/dkv483.
10. Kiehlbauch JA, Plikaytis BD, Swaminathan B, Cameron DN, Wachsmuth IK. 1991. Restriction fragment length polymorphisms in the ribosomal genes for species identification and subtyping of aerotolerant *Campylobacter* species. J Clin Microbiol 29:1670–1676.
11. Vandamme P, Vancanneyt M, Pot B, Mels L, Hoste B, Dewettinck D, Vlaes L, van den Borre C, Higgins R, Hommez J, Kersters K, Butzler J-P, Goossens H. 1992. Polyphasic taxonomic study of the emended genus *Arcobacter* with *Arcobacter butzleri* comb. nov. and *Arcobacter skirrowii* sp. nov., an aerotolerant bacterium isolated from veterinary specimens. Int J Syst Bacteriol 42:344–356. https://doi.org/10.1099/00207713-42-3-344.
12. Kiehlbauch JA, Brenner DJ, Nicholson MA, Baker CN, Patton CM, Steigerwalt AG, Wachsmuth IK. 1991. *Campylobacter butzleri* sp. nov. isolated from humans and animals with diarrheal illness. J Clin Microbiol 29:376–385.
13. 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.
14. Miller WG, Yee E, Bono JL. 2018. Complete genome sequence of the *Arcobacter molluscorum* type strain LMG 25693. Microbiol Resour Announc 7:e01293-18. https://doi.org/10.1128/MRA.01293-18.
15. Pacific Biosciences. 2015. Procedure and checklist: 20 kb template preparation using BluePippin size-selection system. Pacific Biosciences, Menlo Park, CA. https://www.pacb.com/wp-content/uploads/2015/09/Procedure-Checklist-20-kb-Template-Preparation-Using-BluePippin-Size-Selection.pdf.
16. Besemer J, Borodovsky M. 2005. GeneMark: Web software for gene finding in prokaryotes, eukaryotes and viruses. Nucleic Acids Res 33:W451–W454. https://doi.org/10.1093/nar/gki487.
17. Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35:3100–3108. https://doi.org/10.1093/nar/gkm160.
18. Laslett D, Canback B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res 32:11–16. https://doi.org/10.1093/nar/gkh152.