AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequences of the *Campylobacter fetus* subsp. *venerealis*, *Campylobacter lari* subsp. *concheus*, *Campylobacter sputorum* bv. sputorum, and *Campylobacter volucris* Type Strains

William G. Miller,[a] Emma Yee[a]

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

ABSTRACT *Campylobacter* spp. are recovered from a wide variety of sources, including birds, livestock, shellfish, and human clinical samples. We present here the complete genomic data for the type strains of *Campylobacter fetus* subsp. *venerealis*, *Campylobacter lari* subsp. *concheus*, *Campylobacter sputorum* bv. sputorum, and *Campylobacter volucris*.

The genus *Campylobacter* currently contains 37 taxa comprising 31 validly described species, with five of these species further divided into 11 subspecies. Genomic data exist for the type strains of 34 *Campylobacter* taxa (e.g., see references 1–8). Such information would be valuable as baseline reference data for each species. Although genomic data exist for strains of the remaining three taxa (i.e., *Campylobacter lari* subsp. *concheus*, *Campylobacter sputorum*, and *Campylobacter volucris* [7, 9]), no genomic data, draft or complete, exist for their type strains. Additionally, while the type genome of *Campylobacter fetus* subsp. *venerealis* has been sequenced (10), a complete genome (one gapless contig) is recommended due to the presence of both mobile elements and a suite of S-layer genes with repeated motifs in *Campylobacter fetus* subsp. *venerealis* strains (11, 12). In this study, we report the closed genome sequences of *Campylobacter lari* subsp. *concheus* LMG 21009^T, *Campylobacter sputorum* bv. sputorum LMG 7795^T, *Campylobacter volucris* LMG 24380^T, and *Campylobacter fetus* subsp. *venerealis* NCTC 10354^T.

The four type strains were obtained from the NCTC (*C. fetus* subsp. *venerealis*) or BCCM/LMG (*C. lari* subsp. *concheus*, *C. sputorum* bv. sputorum, and *C. volucris*) culture collections. All strains were grown, both initially and for one subculture, microaerobically, at 37°C for 48 h on brain heart infusion agar (Thermo Fisher Scientific, Waltham, MA) amended with 5% laked horse blood. Genomic DNAs (gDNAs) were prepared from loopfuls (~5 μl) of cells using the Promega Wizard genomic DNA purification kit. For each strain, a single preparation of gDNA was used to construct the Illumina and PacBio libraries. Each Illumina library was prepared using the Illumina Nextera DNA Flex library prep kit with 20 ng of gDNA. Pooled libraries were sequenced on a MiSeq instrument at 8.0 pM with dual-index paired-end reads using the MiSeq reagent kit v2 (300 cycles). Each 20-kb PacBio library was prepared using 15 μg of gDNA, the PacBio SMRTbell template prep kit v1.0, and the manufacturer's protocols. Single-molecule real-time (SMRT) sequencing was performed on an RS II sequencer. PacBio reads were assembled using the Hierarchical Genome Assembly Process (HGAP) v3.0 in SMRT Analysis software v2.3.0. Default parameters were used for all software unless otherwise specified. The sequencing metrics are presented in Table 1. SMRT sequencing resulted in a single chromosomal contig for each of the four strains; for two strains (NCTC 10354^T and LMG 24380^T), single plasmid contigs were also present. Almost all of the low-quality (Q < 40)

Citation Miller WG, Yee E. 2019. Complete genome sequences of the *Campylobacter fetus* subsp. *venerealis*, *Campylobacter lari* subsp. *concheus*, *Campylobacter sputorum* bv. sputorum, and *Campylobacter volucris* type strains. Microbiol Resour Announc 8:e01157-19. https://doi.org/10.1128/MRA.01157-19.

Editor David A. Baltrus, University of Arizona

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

Received 17 September 2019
Accepted 14 October 2019
Published 7 November 2019

TABLE 1 Sequencing metrics and genomic data

Feature[a]	Value(s) for[b]:			
	C. fetus subsp. *venerealis* NCTC 10354^T	*C. lari* subsp. *concheus* LMG 21009^T	*C. sputorum* bv. sputorum LMG 7795^T	*C. volucris* LMG 24380^T
Source	Heifer, vaginal	Shellfish	Human, oral	Black-headed gull
Sequencing metrics				
Illumina MiSeq platform				
No. of reads	1,460,738	1,699,532	1,651,458	1,702,578
No. of bases	216,732,966	253,024,312	243,920,520	252,662,467
Avg length (bases)	148	149	148	148
Coverage (×)	115	170	139	166
PacBio platform				
No. of reads	188,065	328,084	45,750	198,092
No. of bases	1,331,563,178	1,582,521,489	246,714,448	1,189,853,205
Avg length (bases)	7,080.3	4,823.5	5,392.7	6,006.6
Coverage (×)	706	1,063	141	780
Accession no.				
GenBank (chromosome)	CP043435	CP043426	CP043427	CP043428
GenBank (plasmid)	CP043436	NA	NA	CP043429
NCBI SRA	SRP219907	SRP219934	SRP219932	SRP219757
Genomic data				
Chromosome				
Size (kbp)[c]	1,885.7	1,488.5	1,753.4	1,524.9
G+C content (%)	33.29	29.77	29.71	28.57
No. of CDS[d]	1,808	1,443	1,662	1,482
Assigned function (% CDS)	915 (50.6)	865 (59.9)	914 (55.0)	873 (58.9)
General function annotation (% CDS)	510 (28.2)	345 (23.9)	430 (25.9)	357 (24.1)
Domain/family annotation only (% CDS)	106 (5.9)	96 (6.7)	99 (6.0)	87 (5.9)
Hypothetical (% CDS)	277 (15.3)	137 (9.5)	219 (13.2)	165 (11.1)
No. of pseudogenes	50	18	52	19
No. of GC tracts ≥8 bp (no. hypervariable)	33 (28)	13 (13)	17 (11)	17 (17)
Plasmid (size [bp])	p3226 (27,915)	No	No	p9726 (7,145)
Genomic island(s)/CRISPR				
No. of genetic islands	5	0	1	1
No. of CDS in genetic islands	165 [9]	0	22 [1]	48 [1]
CRISPR/Cas loci	Type III	0	Type I-B	0
No. of insertion sequence elements	6; 3	0	[1]	0
Gene content/pathways				
Signal transduction				
No. of Che proteins	6	7	9	7
No. of methyl-accepting chemotaxis proteins	10 [5]; 0	11 [2]	10 [5]	13 [1]; 0
No. of RRs	13; 0	9	15	9; 0
No. of HKs	9 [1]; 0	6	12 [2]	6; 0
No. of RR/HK fusions	1; 0	0	0	0
No. of diguanylate cyclases/phosphodiesterases	9; 0	2	7	2; 0
No. of other signal transduction genes	2; 0	2	3 [1]	2; 0
Motility				
Flagellin genes	*flaAB*	*flaAB*	*fla1-2*	*flaAB*
Restriction/modification systems				
No. of type I systems (*hsd*)	1; 0	1	1	[1]; 0
No. of type II systems	1; 0	2	3	3; 0
No. of type III systems	0	0	0	0
Transcription/translation				
No. of transcriptional regulatory proteins	25 [2]; 0	22 [1]	24 [1]	18; 0
Sigma factors	σ^{28}, σ^{54}, σ^{70}	σ^{28}, σ^{54}, σ^{70}	σ^{28}, σ^{54}, σ^{70}	σ^{28}, σ^{54}, σ^{70}
No. of tRNAs, ribosomal loci	44, 3	46, 3	47, 3	46, 3
Catalase gene (*katA*)	Yes	Yes	No	Yes
Cytolethal distending toxin genes (*cdtABC*)	Yes (3×)	Yes	No	Yes
Histidine biosynthesis genes[e]	I	II	I	II
Nuo cluster[f]	I	II	III	II
Tryptophan biosynthesis [*trpABC*(*DG*)*EF*]	Yes	No	Yes	No

[a] RRs, response regulators; HKs, histidine kinases.
[b] Numbers in brackets indicate pseudogenes or fragments. Values following a semicolon are plasmid borne. NA, not applicable.
[c] Size listed in kilobase pairs due to length variation at the hypervariable GC tracts.
[d] Numbers do not include pseudogenes. CDS, coding sequences.
[e] I represents genes *hisA*, *hisB*, *hisC*, *hisD*, *hisF*, *hisG*(S), *hisH*, *hisIE*, *hisJ*, and *hisZ*; II represents genes *hisA*, *hisBJ*, *hisC*, *hisD*, *hisF*, *hisG*(L), *hisH*, and *hisIE*.
[f] I represents *nuoA-nuoB-nuoCD-nuoE-nuoF-nuoG-nuoH-nuoI-nuoJ-nuoK-nuoL-nuoM-nuoN*; II represents genes *nuoA-nuoB-nuoC-nuoD*-ORF-ORF-*nuoG-nuoH-nuoI-nuoJ-nuoK-nuoL-nuoM-nuoN*; and III represents genes *nuoA-nuoB-nuoC-nuoD-nuoE-nuoF-nuoG-nuoH-nuoI-nuoJ-nuoK-nuoL-nuoM-nuoN*.

bases in these contigs were at the contig ends. Therefore, contig circularization removed these low-quality bases; all contigs were circularized manually with Geneious Prime v2019.1.3 (Biomatters Ltd., Auckland, New Zealand). Base calling was further improved by mapping the quality-trimmed MiSeq reads onto the circularized PacBio contigs within Geneious Prime. Using the Geneious "find variations/SNPs" module with a default minimum variation of 0.3, variations in each genome were identified and corrected to the MiSeq consensus sequence. The final coverages ranged from 280 to 1,233×.

The genomic data are presented in Table 1. Protein-, rRNA-, and tRNA-encoding genes were identified using GeneMark, RNAmmer, and ARAGORN (13–15), respectively, and annotated as described previously (16). The genome sizes, G+C content values, and gene contents of these four strains are similar to those reported for other strains of the same taxa (7, 9, 12). With these genome sequences, genomic data are now available for type strains representing all current taxa in the genus *Campylobacter*.

Data availability. The complete genome sequences of the four *Campylobacter* type strains have been deposited in GenBank, and all MiSeq and PacBio sequencing reads have been deposited in the NCBI Sequence Read Archive (SRA). The accession numbers are provided in Table 1.

ACKNOWLEDGMENT

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

REFERENCES

1. Gilbert MJ, Miller WG, Yee E, Blaser MJ, Wagenaar JA, Duim B. 2013. Complete genome sequence of *Campylobacter fetus* subsp. *testudinum* strain 03-427^T. Genome Announc 1:e01002-13. https://doi.org/10.1128/genomeA.01002-13.
2. Gilbert MJ, Miller WG, Yee E, Kik M, Wagenaar JA, Duim B. 2014. Complete genome sequence of *Campylobacter iguaniorum* strain 1485E^T, isolated from a bearded dragon (*Pogona vitticeps*). Genome Announc 2:e00844-14. https://doi.org/10.1128/genomeA.00844-14.
3. Gilbert MJ, Zomer AL, Timmerman AJ, Spaninks MP, Rubio-Garcia A, Rossen JW, Duim B, Wagenaar JA. 2018. *Campylobacter blaseri* sp. nov., isolated from common seals (*Phoca vitulina*). Int J Syst Evol Microbiol 68:1787–1794. https://doi.org/10.1099/ijsem.0.002742.
4. Miller WG, Chapman MH, Yee E, Revez J, Bono JL, Rossi M. 2017. Complete genome sequence of the hippuricase-positive *Campylobacter avium* type strain LMG 24591. Genome Announc 5:e01221-17. https://doi.org/10.1128/genomeA.01221-17.
5. Miller WG, Yee E. 2015. Complete genome sequence of *Campylobacter gracilis* ATCC 33236^T. Genome Announc 3:e01087-15. https://doi.org/10.1128/genomeA.01087-15.
6. Miller WG, Yee E, Bono JL. 2017. Complete genome sequence of the *Campylobacter helveticus* type strain ATCC 51209. Genome Announc 5:e00398-17. https://doi.org/10.1128/genomeA.00398-17.
7. 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.
8. Miller WG, Yee E, Revez J, Bono JL, Rossi M. 2017. Complete genome sequence of the *Campylobacter cuniculorum* type strain LMG 24588. Genome Announc 5:e00543-17. https://doi.org/10.1128/genomeA.00543-17.
9. Miller WG, Yee E, Chapman MH, Bono JL. 2017. Comparative genomics of all three *Campylobacter sputorum* biovars and a novel cattle-associated *C. sputorum* clade. Genome Biol Evol 9:1513–1518. https://doi.org/10.1093/gbe/evx112.
10. Stynen AP, Lage AP, Moore RJ, Rezende AM, de Resende VD, Ruy PDC, Daher N, de Melo Resende D, de Almeida SS, de Castro Soares S, de Abreu VAC, Rocha AACM, dos Santos AR, Barbosa EG, Costa DF, Dorella FA, Miyoshi A, de Lima AR, Campos FD, de Sá PG, Lopes TS, Rodrigues RM, Carneiro AR, Leão T, Cerdeira LT, Ramos RTJ, Silva A, Azevedo V, Ruiz JC. 2011. Complete genome sequence of type strain *Campylobacter fetus* subsp. *venerealis* NCTC 10354^T. J Bacteriol 193:5871–5872. https://doi.org/10.1128/JB.05854-11.
11. van der Graaf-van Bloois L, Miller WG, Yee E, Bono JL, Rijnsburger M, Campero C, Wagenaar JA, Duim B. 2014. First closed genome sequence of *Campylobacter fetus* subsp. *venerealis* bv. intermedius. Genome Announc 2:e01246-13. https://doi.org/10.1128/genomeA.01246-13.
12. van der Graaf-van Bloois L, Miller WG, Yee E, Rijnsburger M, Wagenaar JA, Duim B. 2014. Inconsistency of phenotypic and genomic characteristics of *Campylobacter fetus* subspecies requires reevaluation of current diagnostics. J Clin Microbiol 52:4183–4188. https://doi.org/10.1128/JCM.01837-14.
13. 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.
14. 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.
15. 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.
16. 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.