AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements™

PROKARYOTES

Complete Genome Sequence of *Campylobacter concisus* ATCC 33237^T and Draft Genome Sequences for an Additional Eight Well-Characterized *C. concisus* Strains

Angela J. Cornelius,[a] William G. Miller,[b] Albert J. Lastovica,[c] Stephen L. W. On,[d] Nigel P. French,[e] Olivier Vandenberg,[f,g] Patrick J. Biggs[e,h]

Institute of Environmental Science and Research Ltd. (ESR), Christchurch, New Zealand[a]; U.S. Department of Agriculture, Agricultural Research Service, Albany, California, USA[b]; University of Western Cape, Bellville, South Africa[c]; Lincoln University, Lincoln, New Zealand[d]; Massey University, Palmerston North, New Zealand[e]; National Reference Centre for Campylobacter, Laboratoire Hospitalier Universitaire de Bruxelles (LHUB-ULB), Brussels, Belgium[f]; School of Public Health, Université Libre de Bruxelles, Brussels, Belgium[g]; New Zealand Genomics Limited, Palmerston North, New Zealand[h]

ABSTRACT We report the complete genome sequence of the *Campylobacter concisus* type strain ATCC 33237 and the draft genome sequences of eight additional well-characterized *C. concisus* strains. *C. concisus* has been shown to be a genetically heterogeneous species, and these nine genomes provide valuable information regarding the diversity within this taxon.

Received 6 June 2017 **Accepted** 8 June 2017 **Published** 20 July 2017

Citation Cornelius AJ, Miller WG, Lastovica AJ, On SLW, French NP, Vandenberg O, Biggs PJ. 2017. Complete genome sequence of *Campylobacter concisus* ATCC 33237^T and draft genome sequences for an additional eight well-characterized *C. concisus* strains. Genome Announc 5:e00711-17. https://doi.org/10.1128/genomeA.00711-17.

Address correspondence to Angela J. Cornelius, angela.cornelius@esr.cri.nz.

The cells of *Campylobacter concisus* are Gram-negative, non-spore-forming (1), small (0.5 × 4 μm), and curved with rounded ends (2). *C. concisus* has been isolated from a variety of sites from the human body, including the gingival crevices of patients with gingivitis and periodontitis, stomach and esophagus biopsy specimens, blood, and both normal and diarrheic stools (2). In South Africa, *C. concisus* is the second most commonly isolated *Campylobacter* species in pediatric diarrheic stools (3). This species has also been shown to be phenotypically (4, 5) and genetically (6–11) heterogeneous.

Nine strains were sequenced in this study. *C. concisus* ATCC 33237 is the type strain of this species and was sequenced to completion. One strain, CCUG 19995, was isolated in 1987 in Sweden from a patient with pyrexia and exanthema. The remaining seven strains (Lasto28.99, Lasto61.99, Lasto64.99, Lasto127.99, Lasto205.94, Lasto220.96, and Lasto393.96) were isolated in South Africa between 1994 and 1999 from patients with dysentery, diarrhea, or loose mucoid stools. Strains CCUG 19995, Lasto127.99, and Lasto393.96 are from genomospecies 2, 5, and 6 (12, 13), respectively, while the remaining six strains are members of genomospecies 1 (12). The draft genomes of the eight strains have been well characterized and are genetically diverse (12).

Sequencing of ATCC 33237^T was undertaken using the 454 FLX+ (Titanium chemistry), Illumina (HiSeq), and PacBio platforms. The 454 and Illumina reads were assembled using Newbler (version 2.6) (14, 15) into a single scaffold that was closed using PCR amplification and Sanger sequencing. PacBio sequencing was performed to address repeat regions within the genome and an optical bacterial restriction map (16, 17) (restriction enzyme *Spe*I; OpGen, Gaithersburg, MD) was used to validate the assembly. Protein-coding genes, ribosomal loci, tRNAs, and gene start points were identified as described (18). Annotation was performed by BLASTP comparison to the proteomes of completed *Campylobacter* genomes or to proteins in the NCBI nonredundant database, and by identification of Pfam domains (v.27.0) (19).

Sequencing of CCUG 19995, Lasto28.99, Lasto61.99, Lasto64.99, Lasto127.99, Lasto205.94, Lasto220.96, and Lasto393.96 was undertaken using an Illumina MiSeq. Average coverage between 152× and 254× was achieved. Velvet (version 1.2.10) (20) was used to assemble the short reads, which were quality trimmed using SolexaQA++ (21) at a quality threshold of 0.01, and then sorted by length to remove all resulting reads less than 50 bases long. The draft genomes were annotated using the Prokaryotic Genome Annotation Pipeline (22). The N_{50} values for these genomes, as calculated using the QUAST (23) online calculator (http://quast.bioinf.spbau.ru/), were between 134,605 and 349,534 bp.

The two genomes from genomospecies 2 and 5 (CCUG 19995 and Lasto127.99) had G+C contents of 39.4%, compared to G+C values of between 37.4% and 37.7% for the seven genomes from genomospecies 1 and 6.

Accession number(s). The genome sequences of ATCC 33237^T, CCUG 19995, Lasto28.99, Lasto61.99, Lasto64.99, Lasto127.99, Lasto205.94, Lasto220.96, and Lasto393.96 have been deposited at GenBank under the accession numbers CP012541, NDYN00000000, NDYO00000000, NEFM00000000, NDYP00000000, NDYQ00000000, NDYR00000000, NDYS00000000, and NDYT00000000, respectively. The versions described in this paper are the first versions.

ACKNOWLEDGMENTS

This work was supported by the New Zealand Ministry for Business, Innovation and Employment through ESR Core Funding.

The Illumina HiSeq sequencing was undertaken by SeqWright Genomic Services (GE Healthcare); the PacBio by James L. Bono at USDA, ARS, Clay Center; and the Illumina MiSeq by New Zealand Genomics Limited (NZGL), Massey University.

REFERENCES

1. Vandamme P, Dewhirst FE, Paster BJ, On SLW. 2005. Family I. *Campylobacteraceae*, p 1145–1146. *In* Brenner DJ, Krieg NR, Staley JT (ed), Bergey's manual of systematic bacteriology, 2nd ed, vol 2. The *Proteobacteria*. Part C. The *Alpha-*, *Beta-*, *Delta-*, and *Epsilonproteobacteria*. Springer, New York, NY.
2. Vandamme P, Dewhirst FE, Paster BJ, On SLW. 2005. Genus I. *Campylobacter*, p 1147–1160. *In* Brenner DJ, Krieg NR, Staley JT (ed), Bergey's manual of systematic bacteriology, 2nd ed, vol 2. The *Proteobacteria*. Part C. The *Alpha-*, *Beta-*, *Delta-*, and *Epsilonproteobacteria*. Springer, New York, NY.
3. Lastovica AJ. 2006. Emerging *Campylobacter* spp.: the tip of the iceberg. Clin Microbiol Newsl 28:49–56. https://doi.org/10.1016/j.clinmicnews.2006.03.004.
4. Aabenhus R, Permin H, Andersen LP. 2005. Characterization and subgrouping of *Campylobacter concisus* strains using protein profiles, conventional biochemical testing and antibiotic susceptibility. Eur J Gastroenterol Hepatol 17:1019–1024. https://doi.org/10.1097/00042737-200510000-00003.
5. On SLW, Holmes B, Sackin MJ. 1996. A probability matrix for the identification of campylobacters, helicobacters and allied taxa. J Appl Bacteriol 81:425–432.
6. Aabenhus R, On SL, Siemer BL, Permin H, Andersen LP. 2005. Delineation of *Campylobacter concisus* genomospecies by amplified fragment length polymorphism analysis and correlation of results with clinical data. J Clin Microbiol 43:5091–5096. https://doi.org/10.1128/JCM.43.10.5091-5096.2005.
7. Deshpande NP, Kaakoush NO, Wilkins MR, Mitchell HM. 2013. Comparative genomics of *Campylobacter concisus* isolates reveals genetic diversity and provides insights into disease association. BMC Genomics 14:585. https://doi.org/10.1186/1471-2164-14-585.
8. Ismail Y, Mahendran V, Octavia S, Day AS, Riordan SM, Grimm MC, Lan R, Lemberg D, Tran TA, Zhang L. 2012. Investigation of the enteric pathogenic potential of oral *Campylobacter concisus* strains isolated from patients with inflammatory bowel disease. PLoS One 7:e38217. https://doi.org/10.1371/journal.pone.0038217.
9. Kaakoush NO, Deshpande NP, Wilkins MR, Raftery MJ, Janitz K, Mitchell H. 2011. Comparative analyses of *Campylobacter concisus* strains reveal the genome of the reference strain BAA-1457 is not representative of the species. Gut Pathog 3:15. https://doi.org/10.1186/1757-4749-3-15.
10. Kalischuk LD, Inglis GD. 2011. Comparative genotypic and pathogenic examination of *Campylobacter concisus* isolates from diarrheic and non-diarrheic humans. BMC Microbiol 11:53. https://doi.org/10.1186/1471-2180-11-53.
11. Van Etterijck R, Breynaert J, Revets H, Devreker T, Vandenplas Y, Vandamme P, Lauwers S. 1996. Isolation of *Campylobacter concisus* from feces of children with and without diarrhea. J Clin Microbiol 34:2304–2306.
12. On SLW, Seimer BL, Brandt SM, Chung P, Lastovica AJ. 2013. Characterisation of *Campylobacter concisus* strains from South Africa using amplified fragment length polymorphism (AFLP) profiling and a genomospecies-specific polymerase chain reaction (PCR) assay: identification of novel genomospecies and correlation with clinical data. Afr J Microbiol Res 7:1845–1851. https://doi.org/10.5897/AJMR12.2182.
13. Vandamme P, Falsen E, Pot B, Hoste B, Kersters K, De Ley J. 1989. Identification of EF group 22 campylobacters from gastroenteritis cases as *Campylobacter concisus*. J Clin Microbiol 27:1775–1781.
14. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM. 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature 437:376–380. https://doi.org/10.1038/nature03959.
15. Miller JR, Koren S, Sutton G. 2010. Assembly algorithms for next-generation sequencing data. Genomics 95:315–327. https://doi.org/10.1016/j.ygeno.2010.03.001.
16. Nagarajan N, Read TD, Pop M. 2008. Scaffolding and validation of bacterial genome assemblies using optical restriction maps. Bioinformatics 24:1229–1235. https://doi.org/10.1093/bioinformatics/btn102.
17. Saha S, Rajasekaran S. 2014. Efficient and scalable scaffolding using optical restriction maps. BMC Genomics 15(Suppl 5):S5. https://doi.org/10.1186/1471-2164-15-S5-S5.

18. Merga JY, Winstanley C, Williams NJ, Yee E, Miller WG. 2013. Complete genome sequence of the *Arcobacter butzleri* cattle isolate 7h1h. Genome Announc 1(4):e00655-13. https://doi.org/10.1128/genomeA.00655-13.
19. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer EL, Tate J, Punta M. 2014. Pfam: the protein families database. Nucleic Acids Res 42: D222–D230. https://doi.org/10.1093/nar/gkt1223.
20. Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. Genome Res 18:821–829. https://doi.org/10.1101/gr.074492.107.
21. Cox MP, Peterson DA, Biggs PJ. 2010. SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. BMC Bioinformatics 11:485. https://doi.org/10.1186/1471-2105-11-485.
22. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res 44:6614–6624. https://doi.org/10.1093/nar/gkw569.
23. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086.