CrossMark
click for updates

Complete Genome Sequence of Type Strain *Campylobacter fetus* subsp. *fetus* ATCC 27374

Luciana M. Oliveira,[a,b] Daniela M. Resende,[a,c] Elaine M. S. Dorneles,[d,e] Elvira C. A. Horácio,[a,c] Fernanda L. Alves,[f]
Leilane O. Gonçalves,[a,g] Grace S. Tavares,[a,c] Ana Paula R. Stynen,[d] Andrey P. Lage,[d] Jeronimo C. Ruiz[a,c,g]

Grupo Informática de Biossistemas e Genômica, Centro de Pesquisas René Rachou (CPqRR), Fiocruz Minas, Belo Horizonte, Minas Gerais, Brazil[a]; Programa de Pós-Graduação em Bioinformática (ICB), Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil[b]; Programa de Pós-Graduação em Biologia Computacional e Sistemas, Instituto Oswaldo Cruz (IOC), Fiocruz, Rio de Janeiro, Brazil[c]; Departamento de Medicina Veterinária Preventiva, Laboratório de Bacteriologia Aplicada, Escola de Veterinária, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil[d]; Departamento de Medicina Veterinária, Universidade Federal de Lavras, Lavras, Minas Gerais, Brazil[e]; Programa de Pós-Graduação em Microbiologia (ICB), Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil[f]; Programa de Pós-Graduação em Ciências da Saúde, Centro de Pesquisas René Rachou (CPqRR), Fiocruz Minas, Minas Gerais, Brazil[g]

L.M.O. and D.M.R. contributed equally to this study.

Campylobacter fetus subsp. fetus is a zoonotic bacterium important for animal and public health. The complete sequencing and annotation of the genome of the type strain C. fetus subsp. fetus ATCC 27374 are reported here.

Address correspondence to Jeronimo C. Ruiz, jeronimo@cpqrr.fiocruz.br.

Campylobacter fetus subsp. *fetus* is a zoonotic pathogen that colonizes the intestinal and genital tract of sheep and cattle, causing infertility and abortions (1). In humans, it causes gastroenteritis and occasionally bacteremia and extraintestinal infections (2). The classification of *C. fetus* in subspecies, *C. fetus* subsp. *fetus* or *C. fetus* subsp. *venerealis,* is based on clinical features, host specificity, and phenotypic tests (3).

Nonetheless, a comparison of phenotypic identification and genomic characteristics showed some discordance, especially to *C. fetus* subsp. *fetus* (4). The use of virulence genes to differentiate the *C. fetus* subspecies has been proposed (4–6); however, the small number of *C. fetus* subsp. *fetus* genomes available for comparison precludes the identification of molecular targets with diagnosis potential and a full assessment of virulence factors that determine the specificity of infection. The complete genome of *C. fetus* subsp. *fetus* ATCC 27374 (NCTC 10842, CCUG 6823, CIP 5396, DSM 5361, LMG 6442, Mouton 1), the type strain of the species and subspecies (7), isolated from a sheep fetus brain (8), is reported here.

Sequencing was performed on an Ion Torrent system (Thermo Fisher) with 36.23-fold coverage. A total of 257,181 reads were generated from sequencing, which were adapter and quality trimmed using PRINSEQ (9). Sequences were *de novo* assembled with Velvet (10). The final assembly consists of 96 contigs and N_{50} contig size of 33,054 bp. The assembly is composed of 94 contigs that were aligned and ordered using the ABACAS algorithm (11) and with *C. fetus* subsp. *fetus* 04/554 (accession no. NZ_CP008808) as the reference genome. Two contigs with unknown positions in scaffolds were concatenated to the linear sequence. The open reading frames (ORFs), genes encoding ribosomal RNAs (rRNAs), transfer RNAs (tRNAs), and noncoding RNAs (ncRNAs) were identified with NCBI Prokaryotic Genome Annotation Pipeline version 3.3 (12).

Prediction of transposable elements was performed by TransposonPSI (http://transposonpsi.sourceforge.net/), Tandem Repeat Finder, and BLAST against the Repbase, Dfam, GypsyDB, Pfam, and CDD databases (13–23).

The assembled genome is composed of 1,758,333 bp, with a G+C content of 33.16% and 1,853 ORFs, including four rRNAs, 42 tRNA operons, and three ncRNAs. Gene density is 0.676 genes/kb, and the average gene size is ~889 bases/gene.

Additionally, eight proteins or domains related to transposable elements showing similarity to sequences found in *Campylobacter* spp., *Neisseria meningitidis*, and *Streptomyces albus* were predicted.

The sequenced genome will enhance the understanding of *C. fetus* subsp. *fetus* and differences compared to *C. fetus* subsp. *venerealis*, particularly in genomic plasticity, physiology, and host-pathogen interactions.

Accession number(s). This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession no. MKEI00000000. The version described in this paper is version MKEI01000000.

REFERENCES

1. **Blaser M, Newell D, Thompson S, Zechner E, Nachamkin I, Szymanski C**. 2008. Pathogenesis of *Campylobacter fetus*. Campylobacter:401–428.

2. **Wagenaar JA, van Bergen MA, Blaser MJ, Tauxe RV, Newell DG, van Putten JP**. 2014. *Campylobacter fetus* infections in humans: exposure and disease. Clin Infect Dis **58**:1579–1586. http://dx.doi.org/10.1093/cid/ciu085.

3. **Van Berge M, Van Putten J, Dingle KE, Blaser MJ, Wagenaar J, Nachamkin I, Szymanski C**. 2008. Isolation, identification, subspecies differentiation, and typing of *Campylobacter fetus*, p 213–225. In Nachamkin I, Szymanski C, Blaser MJ (ed), Campylobacter, 3rd ed. ASM Press, Washington, DC.

4. **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. http://dx.doi.org/10.1128/JCM.01837-14.

5. **Moolhuijzen PM, Lew-Tabor AE, Wlodek BM, Agüero FG, Comerci DJ, Ugalde RA, Sanchez DO, Appels R, Bellgard M**. 2009. Genomic analysis of *Campylobacter fetus* subspecies: identification of candidate virulence determinants and diagnostic assay targets. BMC Microbiol **9**:86.

6. **Kienesberger S, Sprenger H, Wolfgruber S, Halwachs B, Thallinger GG, Perez-Perez GI, Blaser MJ, Zechner EL, Gorkiewicz G**. 2014. Comparative genome analysis of *Campylobacter fetus* subspecies revealed horizontally acquired genetic elements important for virulence and niche specificity. PLoS One **9**:e85491. http://dx.doi.org/10.1371/journal.pone.0085491.

7. **Skerman V, McGowan V, Sneath P**. 1980. Approved lists of bacterial names. ASM Press, Washington, DC.

8. **Smith T, Taylor MS**. 1919. Some morphological and biological characters of the spirilla (*Vibrio fetus*, n. sp.) associated with disease of the fetal membranes in cattle. J Exp Med **30**:299–311. http://dx.doi.org/10.1084/jem.30.4.299.

9. **Schmieder R, Edwards R**. 2011. Quality control and preprocessing of metagenomic datasets. Bioinformatics **27**:863–864. http://dx.doi.org/10.1093/bioinformatics/btr026.

10. **Zerbino DR, Birney E**. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. Genome Res **18**:821–829. http://dx.doi.org/10.1101/gr.074492.107.

11. **Assefa S, Keane TM, Otto TD, Newbold C, Berriman M**. 2009. ABACAS: algorithm-based automatic contiguation of assembled sequences. Bioinformatics **25**:1968–1969. http://dx.doi.org/10.1093/bioinformatics/btp347.

12. **Angiuoli SV, Gussman A, Klimke W, Cochrane G, Field D, Garrity G, Kodira CD, Kyrpides N, Madupu R, Markowitz V, Tatusova T, Thomson N, White O**. 2008. Toward an online repository of Standard Operating Procedures (SOPs) for (Meta)genomic annotation. Omics **12**:137–141. http://dx.doi.org/10.1089/omi.2008.0017.

13. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ**. 1990. Basic local alignment search tool. J Mol Biol **215**:403–410. http://dx.doi.org/10.1016/S0022-2836(05)80360-2.

14. **Schäffer AA, Wolf YI, Ponting CP, Koonin EV, Aravind L, Altschul SF**. 1999. IMPALA: matching a protein sequence against a collection of psi-blast-constructed position-specific score matrices. Bioinformatics **15**:1000–1011. http://dx.doi.org/10.1093/bioinformatics/15.12.1000.

15. **Wootton JC, Federhen S**. 1993. Statistics of local complexity in amino acid sequences and sequence databases. Comput Chem **17**:149–163. http://dx.doi.org/10.1016/0097-8485(93)85006-X.

16. **Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A**. 2016. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res **44**:D279–D285. http://dx.doi.org/10.1093/nar/gkv1344.

17. **Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ**. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25**:3389–3402. http://dx.doi.org/10.1093/nar/25.17.3389.

18. **Schäffer AA, Aravind L, Madden TL, Shavirin S, Spouge JL, Wolf YI, Koonin EV, Altschul SF**. 2001. Improving the accuracy of psi-blast protein database searches with composition-based statistics and other refinements. Nucleic Acids Res **29**:2994–3005. http://dx.doi.org/10.1093/nar/29.14.2994.

19. **Marchler-Bauer A, Anderson JB, Cherukuri PF, DeWeese-Scott C, Geer LY, Gwadz M, He S, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Liebert CA, Liu C, Lu F, Marchler GH, Mullokandov M, Shoemaker BA, Simonyan V, Song JS, Thiessen PA**. 2005. CDD: a Conserved Domain Database for protein classification. Nucleic Acids Res **33**:D192–D196. http://dx.doi.org/10.1093/nar/gki069.

20. **Benson G**. 1999. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res **27**:573–580. http://dx.doi.org/10.1093/nar/27.2.573.

21. **Llorens C, Futami R, Covelli L, Dominguez-Escribá L, Viu JM, Tamarit D, Aguilar-Rodriguez J, Vicente-Ripolles M, Fuster G, Bernet GP**. 2011. The Gypsy Database (GyDB) of mobile genetic elements: release 2.0. Nucleic Acids Res **39**:D70–D74.

22. **Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J**. 2005. Repbase update, a database of eukaryotic repetitive elements. Cytogenet Genome Res **110**:462–467. http://dx.doi.org/10.1159/000084979.

23. **Wheeler TJ, Clements J, Eddy SR, Hubley R, Jones TA, Jurka J, Smit AF, Finn RD**. 2013. Dfam: a database of repetitive DNA based on profile hidden Markov models. Nucleic Acids Res **41**:D70–D82. http://dx.doi.org/10.1093/nar/gks1265.