Complete Genome Sequence of *Clostridioides difficile* Ribotype 255 Strain Mta-79, Assembled Using Oxford Nanopore and Illumina Sequencing

Jennifer K. Spinler,[a,b] Anne J. Gonzales-Luna,[c] Sabeen Raza,[a,b] Jessica K. Runge,[a,b] Ruth Ann Luna,[a,b] Tor C. Savidge,[a,b] Kevin W. Garey[c]

[a]Department of Pathology & Immunology, Baylor College of Medicine, Houston, Texas, USA
[b]Texas Children's Microbiome Center, Department of Pathology, Texas Children's Hospital, Houston, Texas, USA
[c]University of Houston College of Pharmacy, Houston, Texas, USA

ABSTRACT Hybrid *de novo* assembly of Illumina/Nanopore sequence data produced a complete circular sequence of the chromosome for a *Clostridioides difficile* ribotype 255 (RT255) isolate from an elderly patient with recurrent *C. difficile* infection (CDI). This provides a high-quality representative sequence for the RT255 lineage.

*C*lostridioides difficile is the leading cause of antibiotic-associated diarrhea and poses an urgent public health threat with >300,000 nosocomial infections in the United States each year (1). Recent epidemiological surveys indicate that *C. difficile* ribotype 255 (RT255) has emerged as the fifth most common RT in Texas and supersedes hypervirulent RTs 027 and 078 in Houston, TX (2). We report the complete genome sequence of an RT255 strain, *C. difficile* Mta-79, to serve as a high-quality reference sequence for the RT255 lineage.

C. difficile Mta-79 was obtained from an 80-year-old female during her third recurrent episode of community-onset health care facility-associated *C. difficile* infection (CDI) in Houston, TX. Cryo-frozen isolates were incubated on cycloserine-cefoxitin fructose agar (CCFA) plates under anaerobic conditions for 48 h. Isolates were confirmed to be *C. difficile* by Gram staining, typical odor, and presence of *C. difficile* antigen on microscreen latex agglutination (Microgen Bioproducts). Genomic DNA was extracted using the MasterPure complete DNA and RNA purification kit (catalog number MC85200; Lucigen). The same DNA prep was used for both Illumina and Nanopore sequencing. For Illumina sequencing, libraries of fragmented genomic DNA were prepared using a NEXTflex rapid DNA-seq kit (catalog number NOVA-5149-02; Bioo Scientific). Paired-end reads (1,117,792 2 × 150-bp reads) were generated on the MiSeq platform (Illumina, San Diego, CA, USA) using the Illumina MiSeq reagent kit v2 (catalog number MS-102-2002) and the PhiX control kit v3 (catalog number FC-110-3001). Raw reads were quality filtered and adapter trimmed with Trimmomatic v0.36 (3). Oxford Nanopore Technologies (ONT) sequencing was carried out on libraries prepared with ONT's rapid barcoding kit (catalog number SQK-RBK004) using a MinION device with flow cell type R9.4.1 (catalog number FLO-MIN106D). Guppy v3.1.5 was used to base call, quality filter (minimum Q score, 10), demultiplex, barcode, and quality trim the 58,759 ONT reads. Multilocus sequence typing analysis of Illumina reads with SRST2 (4) classified *C. difficile* Mta-79 as sequence type 34 (ST34). Nanopore and Illumina reads were combined for hybrid *de novo* assembly using Unicycler v0.4.8-beta (5) in the conservative assembly mode. One circular chromosome of 4,123,685 bp (28.68% G+C content) resulted. The average genome coverage for short- and long-read sequences was 77.3× and 39.5×, respectively, as calculated using BBMap v38.68 (https://sourceforge.net/projects/bbmap/). The assembly was annotated with the NCBI Prokaryotic Annotation

Pipeline (6) and resulted in 3,634 protein-coding genes, 32 rRNA operons, and 9 CRISPR arrays.

Data availability. The genome sequence for *C. difficile* Mta-79 (RT255) has been deposited in NCBI under BioProject PRJNA556848. The Illumina paired-end FASTQ and ONT base-called FASTQ files are available in the Sequence Read Archive under accession numbers SRR9850531 and SRR9850530, respectively. The annotated genome is listed under the NCBI accession number CP042267.

REFERENCES

1. Lessa FC, Mu Y, Bamberg WM, Beldavs ZG, Dumyati GK, Dunn JR, Farley MM, Holzbauer SM, Meek JI, Phipps EC, Wilson LE, Winston LG, Cohen JA, Limbago BM, Fridkin SK, Gerding DN, McDonald LC. 2015. Burden of *Clostridium difficile* infection in the United States. N Engl J Med 372:825–834. https://doi.org/10.1056/NEJMoa1408913.
2. Gonzales-Luna AJ, Carlson TJ, Dotson KM, Lancaster C, Miranda J, Endres BT, Basseres E, Begum K, Alam MJ, Garey KW. 2019. Texas state-wide surveillance of Clostridioides difficile infection, 2011–2018, abstr, p 137. 29th European Congress of Clinical Microbiology and Infectious Diseases, Amsterdam, Netherlands.
3. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
4. Inouye M, Dashnow H, Raven LA, Schultz MB, Pope BJ, Tomita T, Zobel J, Holt KE. 2014. SRST2: rapid genomic surveillance for public health and hospital microbiology labs. Genome Med 6:90. https://doi.org/10.1186/s13073-014-0090-6.
5. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595.
6. 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.