AMERICAN SOCIETY FOR MICROBIOLOGY

Microbiology®
Resource Announcements

Complete Genome Sequence of *Campylobacter jejuni* Strain G1, Isolated from a Patient with Guillain-Barré Syndrome

Andrey V. Karlyshev[a]

aSchool of Life Sciences, Pharmacy and Chemistry, Faculty of Science, Engineering and Computing, Kingston University London, Kingston upon Thames, United Kingdom

ABSTRACT Here, I report the complete genome sequence of *Campylobacter jejuni* strain G1, belonging to Penner serotype HS1. One remarkable feature of the genome of this isolate is the presence of four copies of Mu-like prophages, of which none are present in some other strains, including the reference strain NCTC11168.

Campylobacter jejuni strain G1, isolated from a Guillain-Barré syndrome patient (1), has the smallest known capsular polysaccharide locus (2–6). A draft genome sequence of this strain (7) allowed the identification of genes responsible for the biosynthesis of a highly efficient multidrug efflux pump involved in multidrug resistance (8). In this study, the complete genome sequence of this strain was derived using a hybrid approach (9).

The strain was grown on brain heart infusion agar (Oxoid, USA) for 48 hours at 37°C in a microaerobic atmosphere. Approximately 10^8 CFU were resuspended in 120 μl of Tris-EDTA (TE) buffer with lysozyme (0.1 mg/ml) and RNase A (0.1 mg/ml), incubated for 25 min at 37°C, followed by addition of proteinase K and SDS to 0.1 mg/ml and 0.5%, respectively, and incubated for 5 min at 65°C. Genomic DNA was purified using an equal volume of solid phase reversible immobilization (SPRI) beads (Beckman, USA) and resuspended in elution buffer (Qiagen, Germany).

A sequencing DNA library was prepared using the Nextera XT library prep kit (Illumina, San Diego, CA) following the manufacturer's protocol. Assembly of $2 \times 250,566$ paired short reads (average, 228 bp) produced by the Illumina NovaSeq 6000 platform using SPAdes software v.3.7 resulted in 25 contigs (1 to 406 kb); the N_{50} value was 120,957 bp. Long read genomic DNA libraries were prepared with the Oxford Nanopore Technologies (ONT; United Kingdom) SQK-RBK004 kit and/or SQK-LSK109 kit with the native barcoding EXP-NBD104/114 (ONT) kit with 500 ng of DNA. Barcoded samples were pooled into a single sequencing library and loaded in a FLO-MIN106 (R.9.4.1) flow cell in a GridION system (ONT). Reads were adapter trimmed using Trimmomatic 0.30 with a sliding window quality cutoff of Q15 (10). Genome assembly was performed using Unicycler v.0.4.9b (11). Default parameters were used for all software unless otherwise specified.

The contigs were joined by using the GridION Oxford Nanopore system, which produced 6,413 reads up to 161 kb with an N_{50} value of 40,557 bp. The size of the complete genome sequence was 1,778,460 bp with GC content of 30.43% and 153.96× coverage. It was annotated by the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v.5.1 (12) using the best-placed reference protein set and GeneMarkS-2+, which identified 1,894 genes, including 1,783 protein-coding genes, 55 pseudogenes, 44 tRNAs, 3 sets of rRNAs (5S, 16S, and 23S), and 3 noncoding RNAs (ncRNAs).

A comparison with the draft genome assembly using Mauve software v.2.4.0 (13) revealed misassembled contigs, despite that they were verified by read mapping, confirming previously reported issues when using read mapping for assessment of draft genome assemblies (14). The draft assembly was 56,930 bp smaller than the complete version. The additional sequences in the complete version were identified by the Mauve tool (13) revealing long repeats represented by four Mu-like prophages. The sequences of three
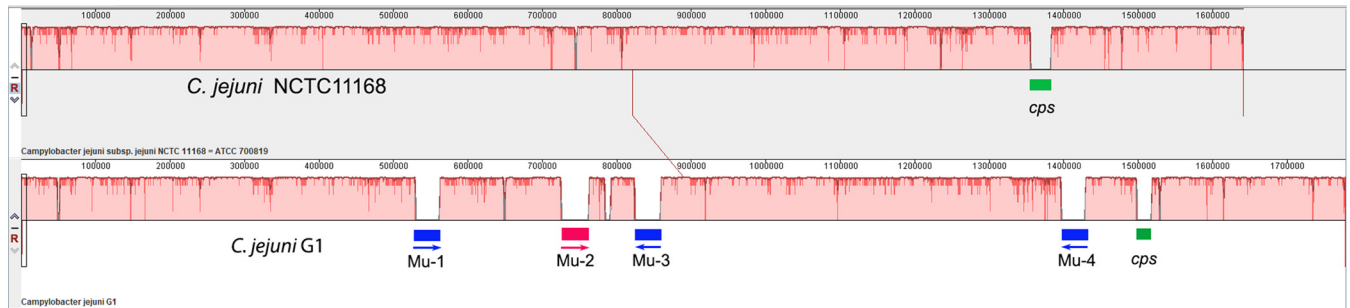
FIG 1 Comparison of complete genomes of *C. jejuni* NCTC11168 and G1 strains using Mauve alignment tool v.2.4.0 (13). Gaps in the genome of strain G1 indicate three almost identical Mu-like prophages (in blue) and a more distantly related Mu-like prophage (in red). Green gaps in both genomes denote capsular polysaccharide loci having very low sequence similarity.

prophages were almost identical (32 kb) and shared extensive similarities with the fourth prophage (37 kb). Mainly due to the presence of these prophages, the genome of G1 was found to be 137 kb larger than that of the reference strain NCTC11168 (15), with the remaining parts of the genomes present in a colinear arrangement (Fig. 1).

The study highlights problems with shotgun draft genome sequencing potentially leading to misassemblies and data misinterpretation and contributes to a better understanding of the evolution of this important pathogen and its lifestyle.

Data availability. This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number CP073712, BioProject number PRJNA261878, and BioSample number SAMN03077634. The version described in this paper is the first version, CP073712.1. Raw sequences were deposited under SRA accession numbers SRR14425908 (long reads) and SRR14425515 (short reads).

ACKNOWLEDGMENT

REFERENCES

1. Gregson N, Rees J, Hughes R. 1997. Reactivity of serum IgG anti-GM1 ganglioside antibodies with the lipopolysaccharide fractions of *Campylobacter jejuni* isolates from patients with Guillain-Barré syndrome (GBS). J Neuroimmunol 73:28–36. https://doi.org/10.1016/S0165-5728(96)00162-2.
2. Karlyshev AV, Linton D, Gregson NA, Lastovica AJ, Wren BW. 2000. Genetic and biochemical evidence of a *Campylobacter jejuni* capsular polysaccharide that accounts for Penner serotype specificity. Mol Microbiol 35:529–541. https://doi.org/10.1046/j.1365-2958.2000.01717.x.
3. Dorrell N, Mangan JA, Laing KG, Hinds J, Linton D, Al-Ghusein H, Barrell BG, Parkhill J, Stoker NG, Karlyshev AV, Butcher PD, Wren BW. 2001. Whole genome comparison of *Campylobacter jejuni* human isolates using a low-cost microarray reveals extensive genetic diversity. Genome Res 11:1706–1715. https://doi.org/10.1101/gr.185801.
4. Karlyshev AV, McCrossan MV, Wren BW. 2001. Demonstration of polysaccharide capsule in *Campylobacter jejuni* using electron microscopy. Infect Immun 69:5921–5924. https://doi.org/10.1128/IAI.69.9.5921-5924.2001.
5. McNally DJ, Lamoureux MP, Karlyshev AV, Fiori LM, Li J, Thacker G, Coleman RA, Khieu NH, Wren BW, Brisson JR, Jarrell HC, Szymanski CM. 2007. Commonality and biosynthesis of the *O*-methyl phosphoramidate capsule modification in *Campylobacter jejuni*. J Biol Chem 282:28566–28576. https://doi.org/10.1074/jbc.M704413200.
6. Karlyshev AV, Wren BW. 2001. Detection and initial characterization of novel capsular polysaccharide among diverse *Campylobacter jejuni* strains using Alcian blue dye. J Clin Microbiol 39:279–284. https://doi.org/10.1128/JCM.39.1.279-284.2001.
7. Vieira A, Ramesh A, Seddon AM, Karlyshev AV. 2017. CmeABC multidrug efflux pump contributes to antibiotic resistance and promotes *Campylobacter jejuni* survival and multiplication in *Acanthamoeba polyphaga*. Appl Environ Microbiol 83:e01600-17. https://doi.org/10.1128/AEM.01600-17.
8. Lehri B, Kukreja K, Vieira A, Zaremba M, Bonney K, Karlyshev AV. 2015. Specific genetic features of *Campylobacter jejuni* strain G1 revealed by genome sequencing. FEMS Microbiol Lett 362:1–3. https://doi.org/10.1093/femsle/fnu064.
9. Lehri B, Seddon AM, Karlyshev AV. 2015. Lactobacillus fermentum 3872 genome sequencing reveals plasmid and chromosomal genes potentially involved in a probiotic activity. FEMS Microbiol Lett 362:fnv068. https://doi.org/10.1093/femsle/fnv068.
10. 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.
11. 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.
12. 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.
13. Darling AE, Mau B, Perna NT. 2010. ProgressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One 5:e11147. https://doi.org/10.1371/journal.pone.0011147.
14. Lehri B, Seddon AM, Karlyshev AV. 2017. The hidden perils of read mapping as a quality assessment tool in genome sequencing. Sci Rep 7:43149. https://doi.org/10.1038/srep43149.
15. Parkhill J, Wren BW, Mungall K, Ketley JM, Churcher C, Basham D, Chillingworth T, Davies RM, Feltwell T, Holroyd S, Jagels K, Karlyshev AV, Moule S, Pallen MJ, Penn CW, Quail MA, Rajandream MA, Rutherford KM, van Vliet AH, Whitehead S, Barrell BG. 2000. The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. Nature 403:665–668. https://doi.org/10.1038/35001088.