The Complete Genome Sequence of *Campylobacter jejuni* Strain S3

Kerry K. Cooper*, Margarethe A. Cooper¹, Andrea Zuccolo², Bibiana Law¹ and Lynn A. Joens¹

¹Department of Veterinary Science and Microbiology, University of Arizona, 1117 E. Lowell St., Tucson, AZ 85721, USA

²Arizona Genomics Institute, School of Plant Sciences, BIO5 Institute, University of Arizona, Tucson, AZ 85721, USA

*Corresponding Author:

Department of Veterinary Science and Microbiology, University of Arizona

1117 E. Lowell St., Tucson, AZ 85721

Telephone: (520) 626-4148

Fax: (520) 621-6366

Email: kcooper@email.arizona.edu
Abstract

*Campylobacter jejuni* is one of the leading causes of bacterial gastroenteritis in the world; however, there is only one complete genome sequence of a poultry strain to date. Here we report the complete genome sequence and annotation of the second poultry strain, *C. jejuni* strain S3. This strain has been shown to be non-motile, a poor invader *in vitro*, and a poor colonizer of poultry after minimal *in vitro* passage.

Key Words

*Campylobacter jejuni*, S3, poultry, genome sequence
Campylobacter jejuni strain S3 is a poultry isolate that was originally cultured from the feces of a chicken (9, 10), and has been shown to be non-invasive to weakly invasive in vitro (11). It is non-motile and was originally described as an effective colonizer of poultry (10, 14); however, after minimal passage in the laboratory, it no longer effectively colonizes poultry.

Using 454 pyrosequencing technology on a GS FLX system (Roche Diagnostics, Branford, CT) with approximately 26-fold coverage of the entire genome, the complete genome sequence for C. jejuni S3 was determined. The initial genome assembly was based on 203,253 reads, and used the Newbler Assembler software which generated a total of 29 large contigs (>500 bp). The initial genome assembly was then compared and complemented using the Celera Assembler software (12) resulting in a total of 11 large contigs (>2500 bp). The large contig sequences were then analyzed for low consensus quality base pairs using the Consed program (6), and any erroneous nucleotides were removed. The large contigs were then organized, aligned and the remaining gaps determined via MUMmer 3.0 software (3) using the genome sequence of C. jejuni RM1221 (13) as a scaffold.

Gap closure was accomplished by primer walking from the large contig sequences and genomic PCR, with the resulting products sequenced using Sanger DNA sequencing on a 3730 DNA Analyzer (Applied Biosystems, Carlsbad, CA). Following construction of a single genomic contig, the sequence was submitted for automatic genome annotation via the RAST server (1). Transfer RNA (tRNA) and ribosomal RNA (rRNA) genes were then annotated through the use of tRNAscan-SE 1.21 (8) and RNAmmer 1.2 (7) programs respectively. Additionally, identified open reading frames (ORF) were confirmed by homology BLAST searches of the NCBI database.
The circular chromosome of *C. jejuni* S3 is composed of 1,681,364 bp with a G+C content of 30.49%, and includes 1,761 putative protein-coding genes or ORFs. Furthermore, the S3 genome contains 28 predicted pseudogenes, 3 rRNA operons and a total of 44 tRNA genes covering all amino acids. *Campylobacter jejuni* S3 also contains a single plasmid that is analogous to the large tetracycline resistance plasmid (pTet) found in *C. jejuni* 81-176 (2). The S3 plasmid has 43,222 bp with a G+C content of 28.99% and contains 49 putative ORFs.

The *C. jejuni* S3 genome contains parts of two of the recently identified *Campylobacter jejuni*-integrated elements (CJIEs) (13), including 28 genes (45.9%) of CJIE1 and 47 genes (82.5%) of CJIE4. *Campylobacter jejuni* S3 CJIE1 also contains 28 genes that are not found in *C. jejuni* RM1221, while CJIE4 from *C. jejuni* S3 is very similar. *C. jejuni* S3 is easier to genetically manipulate (L.A. Joens, unpublished data) than *C. jejuni* RM1221 (4, 5), thus allowing for mutagenesis work with this strain. Overall, the complete genome sequence of *C. jejuni* S3 will be an effective tool for characterizing genes present in CJIE1 and CJIE4, other genes shared by *C. jejuni* RM1221 and *C. jejuni* S3, as well as genes unique to *C. jejuni* S3.
Nucleotide sequence accession number

The complete genome sequence of *C. jejuni* strain S3 is accessible at GenBank under the accession number CP001960, and the entire S3_pTet plasmid sequence can be accessed under the GenBank accession number CP001961.
Acknowledgements

This work was funded by the Arizona Agricultural Experiment Station of the College of Agriculture and Life Sciences at the University of Arizona.


