

# Genomic Variation between *Campylobacter jejuni* Isolates Associated with Milk-Borne-Disease Outbreaks

Joana Revez,<sup>a</sup> Ji Zhang,<sup>a</sup> Thomas Schott,<sup>a,b</sup> Rauni Kivistö,<sup>a</sup> Mirko Rossi,<sup>a</sup> Marja-Liisa Hänninen<sup>a</sup>

Department of Food Hygiene and Environmental Health, Helsinki University, Helsinki, Finland<sup>a</sup>; Biology Oceanography, Leibniz Institute for Baltic Sea Research, Rostock-Warnemünde, Germany<sup>b</sup>

**Bacterial genome sequencing has led to the development of new approaches for the analysis of food-borne epidemics and the exploration of the relatedness of outbreak-associated isolates and their separation from nonassociated isolates. Using Illumina technology, we sequenced a total of six isolates (two from patients, two from raw bulk milk, and two from dairy cattle) associated with a milk-borne *Campylobacter jejuni* outbreak in a farming family and compared their genomes. These isolates had identical pulsed-field gel electrophoresis (PFGE) types, and their multilocus sequence typing (MLST) type was ST-50. We used the Ma\_1 isolate (milk) as the reference, and its genome was assembled and tentatively ordered using the *C. jejuni* NCTC 11168 genome as the scaffold. Using whole-genome MLST (wgMLST), we identified a total of three single-nucleotide polymorphisms (SNPs) and differences in poly(G or C) or poly(A or T) tracts in 12 loci among the isolates. Several new alleles not present in the database were detected. In contrast, the sequences of the unassociated *C. jejuni* strains P14 and 1–12S (both ST-50) differed by 420 to 454 alleles from the epidemic-associated isolates. We found that the fecal contamination of bulk tank milk occurred by highly related sequence variants of *C. jejuni*, which are reflected as SNPs and differences in the length of the poly(A or T) tracts. Poly(G or C) tracts are reversibly variable and are thus unstable markers for comparison. Further, unrelated strains of ST-50 were clearly separated from the outbreak-associated isolates, indicating that wgMLST is an excellent tool for analysis. In addition, other useful data related to the genes and genetic systems of the isolates were obtained.**

*Campylobacter jejuni* is the most common cause of bacterial gastrointestinal infections in the developed world, including Europe. In 2011, a total of 220,209 *Campylobacter* infections were reported in the European Union countries (1). In Finland, a total of 4,200 *Campylobacter* infections were registered in 2013 (see <http://www3.thl.fi/stat>), making it the most common cause of bacterial gastrointestinal infections (70 cases per 100,000 inhabitants). The transmission of this disease is mostly through food or water, but the sources of infections remain mostly unrecognized because most cases are registered from single patients. However, epidemiological studies performed in different countries have shown that eating and handling of chickens are important risk factors for the acquisition of the illness (2, 3). Epidemics associated with drinking unpasteurized milk are frequently reported (4). Small epidemics spreading within farming families caused by drinking raw milk originating from their own farm also occur frequently (5, 6). During the past few years, despite the confirmed scientific data regarding pathogens in raw milk and the public health benefits of pasteurization, certain groups of people concerned with the alleged benefits of the consumption of raw milk and the purported disadvantages of pasteurization have started to drink larger amounts of raw unpasteurized milk (5, 6). An increased number of milk-borne *Campylobacter* outbreaks associated with the controlled commercial sale of raw milk have been reported, e.g., in the United States (7).

A large database (see <http://pubmlst.org/campylobacter>) of multilocus sequence typing (MLST) types of *C. jejuni* isolates from different geographical areas, animal hosts, and human patients has been accumulated, and this database shows the wide genetic diversity among the species found represented as an increasing number of MLST types (8). A recent comparative study on the association of MLST types and whole-genome sequences

(WGS) revealed that the MLST type predicts rather well the genome-level relatedness of *C. jejuni* isolates (9).

Genomic epidemiology has been increasingly applied to the study of epidemic-associated clonal variants of infectious agents to separate them from other related isolates (10, 11). This field combines traditional epidemiological methods with genome sequence similarity analysis of bacterial isolates from patients and their potential reservoirs associated with different times and places. The use of next-generation sequencing (NGS) technology for WGS allows for the sequencing of large numbers of isolates, and novel bioinformatics tools can be used for comparative genomics and analysis of the phylogeny of the isolates. This approach was extensively applied, e.g., during the large multistate *Escherichia coli* O104:H4 epidemic (12). WGS will change the way we compare isolates because it allows for the analysis of all potential diversities and enables the detection of novelties associated with genome synteny, sequences, single genes, and polymorphisms between the isolates, as well as potential changes in the genomes that occur during infection (10, 13). We recently found that only a few single-nucleotide polymorphisms (SNPs) are detected in *C. jejuni* after its intestinal passage through an infected

Received 6 April 2014 Returned for modification 5 May 2014

Accepted 13 May 2014

Published ahead of print 21 May 2014

Editor: W. M. Dunne, Jr.

Address correspondence to Marja-Liisa Hänninen, [marja-liisa.hanninen@helsinki.fi](mailto:marja-liisa.hanninen@helsinki.fi).

Supplemental material for this article may be found at <http://dx.doi.org/10.1128/JCM.00931-14>.

Copyright © 2014, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JCM.00931-14

patient (13). However, there is limited knowledge regarding how to interpret the results of WGS and comparative genomics of outbreak-associated *C. jejuni* isolates and how to combine them with other epidemiological data (i.e., differences in SNPs and homopolymeric tracts in contingency genes).

The aim of the present study was to apply WGS and comparative genomics to six *C. jejuni* milk-borne outbreak-associated isolates (6) in order to determine the sequence-level similarities of the epidemiologically and genotypically linked isolates from patients, raw milk (which has been shown to be the source of the infection), and two dairy cows suspected to be the source of the fecal contamination of the raw milk. Our hypothesis was that the comparison of the genomes of highly related and unrelated *C. jejuni* isolates will produce data that can be used for the interpretation of the results of future genomic epidemiology studies.

## MATERIALS AND METHODS

**Bacterial strains and DNA isolation.** A total of six milk-borne outbreak-associated *C. jejuni* isolates were selected for whole-genome sequencing. These were obtained from two patients (Po\_1 and Po\_2) on 16 December 2002, two bulk milk tank samples from the farm (Ma\_1 and Ma\_B) on 16 December 2002 and 8 January 2003, and two dairy cows from the farm (Le\_204R and Le\_755) on 18 February 2003, as described previously (6). The isolates were retrieved during a long-lasting milk-borne outbreak that started in August 2002 and lasted until December 2002, during which the outbreak affected a total of six members of a farming family. All of the isolates had identical pulsed-field gel electrophoresis (PFGE) profiles with both SmaI and KpnI restriction enzymes, and their MLST type was ST-50 (ST-21 clonal complex) (6). A hen *C. jejuni* 1\_12S strain was used as an unrelated control strain (ST-50). The genomic DNA for genome sequencing was isolated using the Wizard genomic DNA purification kit (Promega, Germany), following the manufacturer's instructions.

**Genome sequencing, assembly, and annotation.** The genome sequence of *C. jejuni* strain Ma\_1 was obtained using Illumina sequencing technology with 100 cycles of paired-end reads, whereas the genomic sequences of the five remaining isolates and the strain 1\_12S were obtained by Illumina HiSeq single-end reads. The assembly and annotation were performed as previously described (13). Briefly, the Illumina reads were trimmed using the ConDeTri Perl script (14) with the default settings and a minimum read length of 75 nucleotides. All of the reads were assembled separately using ABySS (15). Additionally, the contigs of *C. jejuni* Ma\_1 were reordered using Mauve analysis with *C. jejuni* NCTC 11168 as the reference genome. The primary annotation was performed through rapid annotation using subsystem technology (RAST) (16), and the later sequences were manually curated using Artemis. The prophages/integrated elements were searched for in the genomes using IslandViewer (17) and manually inspected.

**Whole-genome MLST.** The reference allele sequences of 1,738 loci for whole-genome MLST (wgMLST) analysis were downloaded from the PubMLST website (see <http://pubmlst.org/campylobacter>) (8, 11). A new allele number was assigned to the sequence if it had at least a 1-bp difference compared with any sequences in the downloaded reference allele sequences, as determined through BLAST+ 2.2.24 (18). All of the truncated and missing loci were excluded from the analysis. The allelic profile data of the commonly shared loci were transformed to a distance matrix (11), and the phylogenetic networks of the isolates were computed based on the Neighbor-Net algorithm using SplitsTree4 (19). As unrelated control genomes for ST-50, *C. jejuni* strain PT14 (20) and a Finnish hen strain, 1\_12S, were used.

**Verification of SNPs and A/T tract differences.** We used PCR and Sanger sequencing for the confirmation of certain allele differences. The primers used for amplification and sequencing can be found in Table S1 in the supplemental material.

**Genome sequence accession numbers.** The genome sequences of *C. jejuni* strain Ma\_1 and the other 6 isolates were submitted to the ENA database under the study accession number PRJEB5994. Their sample accession numbers are ERS452492 (1\_12S), ERS452491 (Po\_2), ERS452490 (Po\_1), ERS452489 (Ma\_B), ERS452449 (Le\_755), ERS452426 (Le\_204R), and ERS452422 (Ma\_1).

## RESULTS

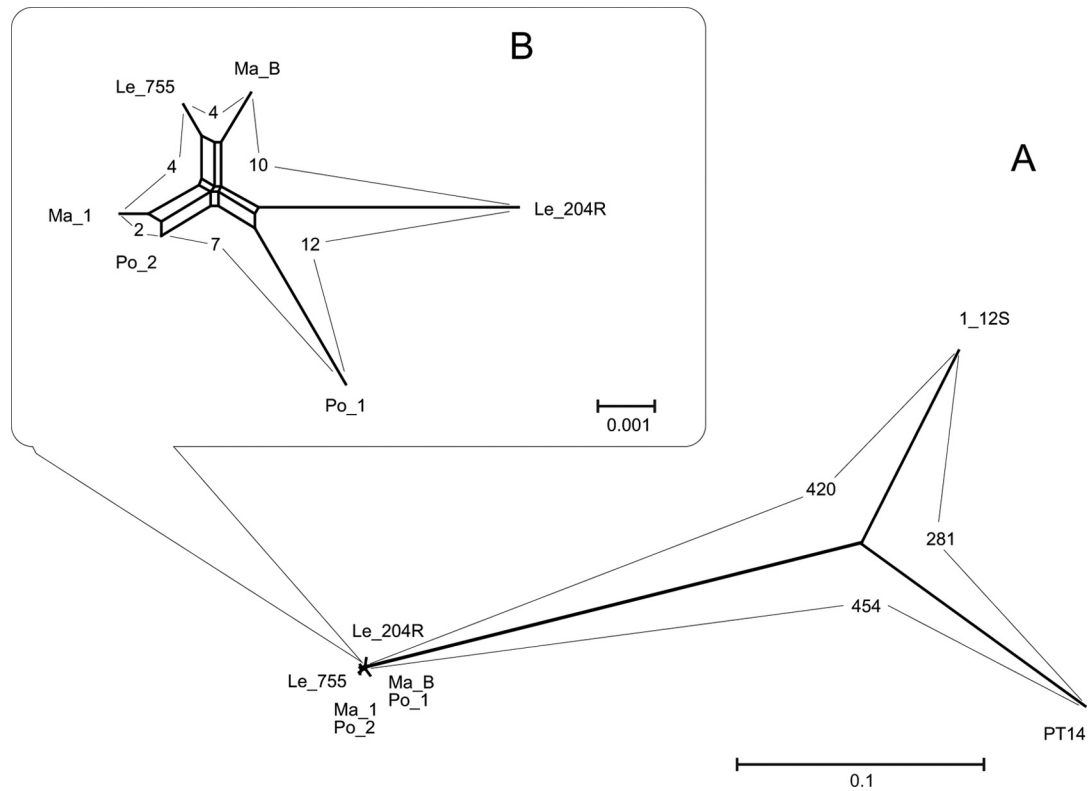
**Genomic features of *C. jejuni* strain Ma\_1.** The sequence data were assembled into 58 contigs, which were later tentatively ordered using the *C. jejuni* NCTC 11168 genome as the scaffold. The assembled genome is composed of a circular chromosome of 1,653,051 bp and includes 1,662 putative protein-coding genes or open reading frames (ORFs). No plasmids or complete integrated elements were detected, as in *C. jejuni* strain RM1221 (21).

**General features of the outbreak-associated isolates.** The six milk-borne outbreak-associated isolates studied were susceptible to both erythromycin and ciprofloxacin (6) but resistant to tetracycline (MIC, 32 mg/liter). All of the isolates had the *tet(O)* gene (22) located in the chromosome inserted into the homolog of Cj0733, thereby disrupting the sequence. All of the outbreak-associated isolates harbored an identical CRISPR (clustered regularly interspaced short palindromic repeats)-Cas system with four repeat regions and three spacers. The CRISPR-associated gene *csn1* was fragmented in Ma\_1 but intact in PT14 (20).

**Whole-genome MLST.** Among the 1,738 studied loci, 1,404 loci were shared between the six outbreak-associated isolates and the control strains PT14 and 1\_12S. The number of allele differences between the outbreak-associated isolates and the non-outbreak-associated strains was in the range of 420 to 454 (Fig. 1; see Table S2 in the supplemental material). The six outbreak-associated isolates shared 1,432 common loci, and 1,417 of these were identical (see Table S3 in the supplemental material). Several new alleles that are not present among the alleles in the PubMLST database were found among the isolates (Table 1). A total of 15 loci exhibited differences among the isolates; eight of these were different due to the changes in the homopolymeric tract lengths (G or C), five were associated with A or T tract length differences, and three were different due to point mutations (SNPs) (Table 1; see Table S3 in the supplemental material). The closest isolates were the milk (Ma\_1) and the patient Po\_2 isolates (two allele differences). Cattle fecal isolate Le\_755 differed by four alleles from Ma\_1. The maximum 12-locus difference was found between Po\_1 and Le\_204R (Fig. 1; see Table S3). Three SNPs and five A or T tract differences were confirmed by PCR and sequence analysis.

## DISCUSSION

*C. jejuni* strains belonging to the ST-21 complex are regarded as generalists because they can colonize a multiplicity of hosts and because they are one of the most common clonal complexes that infect humans (2, 11, 23). A previous genomic study of ST-21 strains (the same clonal complex as ST-50) revealed an overall conservation of the genome synteny for this clonal complex, even though some interstrain diversity was observed (23). In a recent study, Cody et al. (11) compared WGS data from 10 human patient ST-50 isolates and found that most of them were highly divergent ( $\geq 300$  loci were different) and that some strains were highly related (differed by six loci). To reveal the genomic relatedness of the milk-borne outbreak-associated isolates, all six genomes were compared by Neighbor-Net using the genomes of *C.*



**FIG 1** (A) Neighbor-Net algorithm used for constructing phylogenetic networks generated with the 1,404 shared loci of the six milk-borne *C. jejuni* outbreak-associated isolates and two unrelated control strains (*C. jejuni* PT14 and 1\_12S). (B) Neighbor-Net phylogeny generated with the 1,432 shared loci of the six milk-borne isolates. The boldfaced numbers indicate the distances (numbers of allelic differences) between the isolates. For the pairwise comparison of the isolates, see Tables S2 and S3 in the supplemental material.

*jejuni* PT14 (ST-50) and the Finnish hen isolate 1\_12S (ST-50) as nonassociated control strains for the outbreak-associated isolates. This analysis showed that the epidemic-associated isolates differed by a maximum of 12 loci from each other but by several hundreds

of alleles from the two unrelated strains of the same ST. We further found that several alleles of our *C. jejuni* isolates were novel (i.e., not previously present in the wgMLST database), even though this database contains WGS data from a total of 172 ST-50 isolates (see

**TABLE 1** Genomic location, proposed function (homologs of NCTC 11168), and alleles of the 15 polymorphic loci of 1,432 shared loci of the six milk-borne *C. jejuni* outbreak-associated isolates using whole-genome MLST

Locus	No. of alleles in isolate <sup>a</sup> :						Function	Mutation(s)
	Le_204R	Le_755	Ma_1	Ma_B	Po_1	Po_2		
CAMP0019 (Cj0019c)	*349 <sup>+1</sup>	2	*2	2	*349 <sup>+1</sup>	2	MCP <sup>b</sup> -domain signal transduction protein	8 <sup>c</sup> A/9A
CAMP0031 (Cj0031)	241	242	242	241	241	241	Type IIS restriction/modification enzyme	8G/9G
CAMP0157 (Cj0170)	1	1	10	1	2	10	Hypothetical protein	8G/9G/10G
CAMP0608 (Cj0651)	*98 <sup>+1</sup>	1	*1	1	1	1	Integral membrane protein	9T/10T
CAMP0624 (Cj0677)	*107	108	*108	108	*108	108	Potassium-transporting ATPase subunit B	C/T
CAMP0631 (Cj0685c)	7	4	7	4	4	7	Invasion protein CipA	9C/10C
CAMP0935 (Cj1012c)	*115 <sup>+1</sup>	6	*6	6	*6	6	Hypothetical protein	3A/6A
CAMP1031 (Cj1110c)	*239 <sup>+1</sup>	1	1	1	1	1	MCP-type signal transduction protein	6A/7A
CAMP1108 (Cj1189c)	*1	212	*212	212	1	212	Bipartite energy taxis response protein CetB	C/T
CAMP1113 (Cj1194)	*22	22	*22	22	331	22	Phosphate permease	A/G
CAMP1214 (Cj1296)	3	3	4	3	4	4	Hypothetical protein	9G/10G
CAMP1242 (Cj1324)	*27	27	*27	27	*172 <sup>+1</sup>	27	Hypothetical protein	5A/6A
CAMP1243 (Cj1325)	63 <sup>+1</sup>	63 <sup>+2</sup>	63 <sup>+1</sup>	63 <sup>+3</sup>	63 <sup>+2</sup>	63 <sup>+2</sup>	Methyltransferase	9G/10G/11G
CAMP1258 (Cj1342c)	52	273 <sup>+1</sup>	273 <sup>+1</sup>	273 <sup>+1</sup>	273 <sup>+1</sup>	273 <sup>+1</sup>	Motility accessory factor <i>maf7</i>	9G/10G
CAMP1331 (Cj1420c)	1	1	1	70	73	1	Methyltransferase	8G/9G/10G

<sup>a</sup> Superscript numbers on alleles indicate changes that were not listed in the PubMLST database. Asterisks indicate genetic differences that were confirmed by Sanger sequencing before the allelic numbers.

<sup>b</sup> MCP, methyl-accepting chemotaxis protein.

<sup>c</sup> Number of the repeats.

<http://pubmlst.org/campylobacter>). Most of the isolates included in the database were obtained from human patients in Oxford, United Kingdom (11). Further studies are warranted to determine whether these allele differences are explained by the geography-associated microevolution of *C. jejuni*.

We also analyzed which genetic changes explain the allele differences between the isolates. Only three of the differences were caused by mutations in three loci, and 12 were associated with changes in the length of homopolymeric tracts (G or C and A or T). For example, Po\_2 and Ma\_1 had differences in only two loci [both were poly(G) tracts], and Ma\_1 and Ma\_B differed by three loci [all poly(G) tracts]. In addition, Le\_755 was more similar to Ma\_1 than to Le\_204R, and Le\_755 differed from Le\_204R by nine alleles [two SNPs and seven poly(G or C) or poly(A or T) tract lengths]. The G/C homopolymeric tracts in *C. jejuni* are located in contingency genes, and these genes are often phase variable. Furthermore, phase-variable genes are usually located in regions encoding bacterial surface-associated structures, mainly lipooligosaccharide (LOS), flagella, and capsule regions (24). Moreover, all of the outbreak-associated isolates presented variation among the loci from these regions. For example, homologs of the loci CAMP1214, CAMP1242, CAMP1243, and CAMP1258 are located in NCTC 11168 in the flagella region. In our previous study on accidental *C. jejuni* infection, we also found homopolymeric tract changes in some of these loci after passage through a human patient (13). The homolog of CAMP1331 (Cj1420) is a methyltransferase located in the capsule locus. CAMP0157 is a homolog of Cj0170, and this gene was recently shown to be phase variable and associated with motility and colonization success in mice (25). CAMP0031 is a homolog of Cj0031, which is also known to be a contingency gene in *C. jejuni*. Similarly, CAMP0631 is a homolog of Cj0685c, which regulates the surface-associated invasion gene *cipA* (26), and CAMP0935 (a homolog of Cj1012c) encodes a hypothetical membrane protein. SNPs were found in three loci that regulate potassium transport, phosphate permease, and energy taxis. SNP differences were identified between the fecal isolates from two dairy cattle. It appears highly evident that the mutation-associated diversity detected among the outbreak-associated isolates was developed in cattle and may be explained by the rather persistent and long-lasting colonization of dairy cattle by *C. jejuni* (27), which supports opportunities for the accumulation of mutations.

Our results reveal a high level of genome sequence similarity (based on 1,432 shared loci) among the outbreak-associated isolates from the feces of dairy cattle, contaminated milk, and patients, as assessed by wgMLST. Our findings are concordant with the results of other studies performed on *C. jejuni* isolates from infected patients and analyzed over a short period (11, 13). For example, Cody et al. (11) showed that two *C. jejuni* isolates from a patient that were cultivated in the morning and evening on the same day had differences in six loci. Furthermore, through a previous study of the genomes of *C. jejuni* from an infected patient, we found only a few SNPs after human passage (13). In addition, previous experimental animal infection studies found only a few SNPs and changes in the homopolymeric tracts (28), further confirming that the genomic data of the outbreak-related *C. jejuni* isolates represent diversity originating from the original contamination source.

The observations reported by Cody et al. (11) suggest a criterion of a 20-locus difference as a cutoff value, below which

possible clusters can be investigated for the study of clinically and epidemiologically related *C. jejuni* isolates without detailed knowledge of their epidemiological association. However, their analysis of allele differences between highly related isolates in their clusters 1 and 2 also showed only three or four polymorphisms, respectively (11), which is in agreement with our results on *C. jejuni* isolates that are known to originate from an epidemic cluster.

As shown in our results, the analytical tool of the Bacterial Isolate Genome Sequence Database (BIGSdb) accounts for both SNPs and differences in poly(G or C) and poly(A or T) tracts. A total of five loci that were identified as different between the isolates contained A-tract length variations, which is not well characterized in *C. jejuni*. All but one (CAMP0935) of the A/T tracks resulted in truncation of the genes. The poly(G) tracts within the reading frames of the genes in *C. jejuni* are unstable and lead to frequent, stochastic, and reversible changes in the length of repeats (24). The numbers of contingency genes among the different *C. jejuni* strains that have been analyzed to date appear to vary from 29 in NCTC 11168 to 17 in NCTC 81116 (24). Some novel contingency genes may be detected after additional genomic data become available. When we just consider SNPs as the stable differences between the isolates, the studied isolates differ in only three loci.

In conclusion, the real-time analysis of WGS data produced through NGS technology from epidemic-associated *C. jejuni* isolates requires common data analysis tools, such as those available at the PubMLST website (<http://pubmlst.org/campylobacter>), as well as a reference sequence(s) of an unrelated isolate(s) from the same ST. The identification of polymorphic loci requires the differentiation of SNPs and changes associated with differing lengths in G/C homopolymeric tracts. The outbreak-associated isolates appear to differ by only a few SNPs. Further collection of data on outbreaks is required to establish a criterion for the cutoff value for locus differences and for the assessment of the role of homopolymeric tracts in outbreak-associated isolates. Our findings on the similarity of the location of the *tet(O)* gene and the identity of the CRISPR system genes among the outbreak-associated isolates further support the advantages of genome-level analysis, which also results in the accumulation of data on genes and genetic systems among the isolates that will improve our understanding of the ecology of the isolates.

## ACKNOWLEDGMENTS

We thank Urszula Hirvi for her excellent technical assistance.

This work was supported by an Academy of Finland grant on behalf of the Finnish Centre of Excellence in Microbial Food Safety Research (11411405) and the Walter Ehrström Foundation (4703496).

J.Z. was funded by Suomen Kulttuurirahasto (The Finnish Cultural Foundation).

## REFERENCES

1. European Food Safety Authority. 2013. The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2011. EFSA J. 11:3129. <http://dx.doi.org/10.2903/j.efsa.2013.3129>.
2. de Haan CP, Llerena AK, Revez J, Hanninen ML. 2012. Association of *Campylobacter jejuni* metabolic traits with multilocus sequence types. Appl. Environ. Microbiol. 78:5550–5554. <http://dx.doi.org/10.1128/AEM.01023-12>.
3. Sheppard SK, Dallas JF, Strachan NJ, MacRae M, McCarthy ND, Wilson DJ, Gormley FJ, Falush D, Ogdén ID, Maiden MC, Forbes KJ.

2009. *Campylobacter* genotyping to determine the source of human infection. Clin. Infect. Dis. 48:1072–1078. <http://dx.doi.org/10.1086/597402>.
4. Olson CKE, Ethelberg S, van Pelt W, Tauxe RV. 2008. Epidemiology of *Campylobacter jejuni* infections in industrialized nations, p 163–189. In Irving Nachamkin CMS, Martin Blaser J (ed), *Campylobacter*, 3rd ed. American Society for Microbiology, Washington, DC.
  5. Angulo FJ, Lejeune JT, Rajala-Schultz PJ. 2009. Unpasteurized milk: a continued public health threat. Clin. Infect. Dis. 48:93–100. <http://dx.doi.org/10.1086/595007>.
  6. Schildt M, Savolainen S, Hanninen ML. 2006. Long-lasting *Campylobacter jejuni* contamination of milk associated with gastrointestinal illness in a farming family. Epidemiol. Infect. 134:401–405. <http://dx.doi.org/10.1017/S0950268805005029>.
  7. Gould LH, Walsh KA, Vieira AR, Herman K, Williams IT, Hall AJ, Cole D, Centers for Disease Control and Prevention. 2013. Surveillance for foodborne disease outbreaks—United States, 1998–2008. MMWR Surveill. Summ. 62:1–34. <http://dx.doi.org/10.3201/eid1908.121511>.
  8. Jolley KA, Maiden MC. 2010. BIGSdb: scalable analysis of bacterial genome variation at the population level. BMC Bioinformatics 11:595. <http://dx.doi.org/10.1186/1471-2105-11-595>.
  9. Carrillo CD, Kruczkiewicz P, Mutschall S, Tudor A, Clark C, Taboada EN. 2012. A framework for assessing the concordance of molecular typing methods and the true strain phylogeny of *Campylobacter jejuni* and *C. coli* using draft genome sequence data. Front. Cell Infect. Microbiol. 2:57. <http://dx.doi.org/10.3389/fcimb.2012.00057>.
  10. Robinson ER, Walker TM, Pallen MJ. 2013. Genomics and outbreak investigation: from sequence to consequence. Genome Med. 5:36. <http://dx.doi.org/10.1186/gm440>.
  11. Cody AJ, McCarthy ND, Jansen van Rensburg M, Isinkaye T, Bentley SD, Parkhill J, Dingle KE, Bowler IC, Jolley KA, Maiden MC. 2013. Real-time genomic epidemiological evaluation of human *Campylobacter* isolates by use of whole-genome multilocus sequence typing. J. Clin. Microbiol. 51:2526–2534. <http://dx.doi.org/10.1128/JCM.00066-13>.
  12. Rohde H, Qin J, Cui Y, Li D, Loman NJ, Hentschke M, Chen W, Pu F, Peng Y, Li J, Xi F, Li S, Li Y, Zhang Z, Yang X, Zhao M, Wang P, Guan Y, Cen Z, Zhao X, Christner M, Kobbe R, Loos S, Oh J, Yang L, Danchin A, Gao GF, Song Y, Li Y, Yang H, Wang J, Xu J, Pallen MJ, Wang J, Aepfelbacher M, Yang R, Consortium EcOHGAC-S. 2011. Open-source genomic analysis of Shiga-toxin-producing *E. coli* O104:H4. N. Engl. J. Med. 365:718–724. <http://dx.doi.org/10.1056/NEJMoa1107643>.
  13. Revez J, Schott T, Llerena AK, Rossi M, Hanninen ML. 2013. Genetic heterogeneity of *Campylobacter jejuni* NCTC 11168 upon human infection. Infect. Genet. Evol. 16:305–309. <http://dx.doi.org/10.1016/j.meegid.2013.03.009>.
  14. Smeds L, Kunstner A. 2011. ConDeTri—a content dependent read trimmer for Illumina data. PLoS One 6:e26314. <http://dx.doi.org/10.1371/journal.pone.0026314>.
  15. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, Birol I. 2009. ABySS: a parallel assembler for short read sequence data. Genome Res. 19:1117–1123. <http://dx.doi.org/10.1101/gr.089532.108>.
  16. Langille MG, Brinkman FS. 2009. IslandViewer: an integrated interface for computational identification and visualization of genomic islands. Bioinformatics 25:664–665. <http://dx.doi.org/10.1093/bioinformatics/btp030>.
  17. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Res. 42:D206–D214. <http://dx.doi.org/10.1093/nar/gkt1226>.
  18. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. BMC Bioinformatics 10:421. <http://dx.doi.org/10.1186/1471-2105-10-421>.
  19. Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23:254–267. <http://dx.doi.org/10.1093/molbev/msj030>.
  20. Brathwaite KJ, Siringan P, Moreton J, Wilson R, Connerton IF. 2013. Complete genome sequence of universal bacteriophage host strain *Campylobacter jejuni* subsp. *jejuni* PT14. Genome Announc. 1:e00969–13. <http://dx.doi.org/10.1128/genomeA.00969-13>.
  21. Fouts DE, Mongodin EF, Mandrell RE, Miller WG, Rasko DA, Ravel J, Brinkac LM, DeBoy RT, Parker CT, Daugherty SC, Dodson RJ, Durkin AS, Madupu R, Sullivan SA, Shetty JU, Ayodeji MA, Shvartsbeyn A, Schatz MC, Badger JH, Fraser CM, Nelson KE. 2005. Major structural differences and novel potential virulence mechanisms from the genomes of multiple *Campylobacter* species. PLoS Biol. 3:e15. <http://dx.doi.org/10.1371/journal.pbio.0030015>.
  22. Dasti JI, Gross U, Pohl S, Lugert R, Weig M, Schmidt-Ott R. 2007. Role of the plasmid-encoded *tet(O)* gene in tetracycline-resistant clinical isolates of *Campylobacter jejuni* and *Campylobacter coli*. J. Med. Microbiol. 56:833–838. <http://dx.doi.org/10.1099/jmm.0.47103-0>.
  23. Gripp E, Hlahla D, Didelot X, Kops F, Maurischat S, Tedin K, Alter T, Ellerbroek L, Schreiber K, Schomburg D, Janssen T, Bartholomaeus P, Hofreuter D, Woltemate S, Uhr M, Brenneke B, Gruning P, Gerlach G, Wieler L, Suerbaum S, Josenhans C. 2011. Closely related *Campylobacter jejuni* strains from different sources reveal a generalist rather than a specialist lifestyle. BMC Genomics 12:584. <http://dx.doi.org/10.1186/1471-2164-12-584>.
  24. Aidley J, Bayliss CD. 2014. Repetitive DNA: a major source of genetic diversity in *Campylobacter* populations? p 55–72. In Sheppard SK, Méric G (ed), *Campylobacter* ecology and evolution. Caister Academic Press, Swansea, United Kingdom.
  25. Artymovich K, Kim JS, Linz JE, Hall DF, Kelley LE, Kalbach HL, Kathariou S, Gaymer J, Paschke B. 2013. A “successful allele” at *Campylobacter jejuni* contingency locus Cj0170 regulates motility; “successful alleles” at locus Cj0045 are strongly associated with mouse colonization. Food Microbiol. 34:425–430. <http://dx.doi.org/10.1016/j.fm.2013.01.007>.
  26. Poli VF, Thorsen L, Olesen I, Wik MT, Jespersen L. 2012. Differentiation of the virulence potential of *Campylobacter jejuni* strains by use of gene transcription analysis and a Caco-2 assay. Int. J. Food Microbiol. 155:60–68. <http://dx.doi.org/10.1016/j.ijfoodmicro.2012.01.019>.
  27. Hakkinen M, Heiska H, Hanninen ML. 2007. Prevalence of *Campylobacter* spp. in cattle in Finland and antimicrobial susceptibilities of bovine *Campylobacter jejuni* strains. Appl. Environ. Microbiol. 73:3232–3238. <http://dx.doi.org/10.1128/AEM.02579-06>.
  28. Wilson DL, Rathinam VA, Qi W, Wick LM, Landgraf J, Bell JA, Plovovich-Jones A, Parrish J, Finley RL, Mansfield LS, Linz JE. 2010. Genetic diversity in *Campylobacter jejuni* is associated with differential colonization of broiler chickens and C57BL/6J IL10-deficient mice. Microbiology 156:2046–2057. <http://dx.doi.org/10.1099/mic.0.035717-0>.