

Original Article

PCR-RFLP Provides Discrimination for Total *flaA* Sequence Analysis in Clinical *Campylobacter jejuni* Isolates

Mahdi Ghorbanalizadgan^{1,2}, Bita Bakhshi^{1*}, and ShahinNajar-Peerayeh¹

¹Department of Bacteriology, Faculty of Medical Sciences, Tarbiat Modares University, Tehran; and

²Department of Bacteriology, Faculty of Medical Sciences, Baqiyatallah University, Tehran, Iran

SUMMARY: The aims of this study were to determine the genetic relatedness among 20 clinical *Campylobacter jejuni* samples isolated from children with diarrhea in Iran and to introduce the best method of discrimination based on flagellin gene (*flaA*) sequence divergence. A total of 400 stool specimens were obtained from children under 5 years of age from July 2012 to June 2013. Primers were designed based on conserved sequences flanking the *flaA* gene that encompassed and amplified the entire *flaA* gene and followed by sequencing and data analysis with MEGA version 6.0.6 software. Ninety amino acids and 560 nucleotide polymorphic sequences were detected within 1,681 bp of the *flaA* sequence of which 43 (2.5%) and 12 (0.7%) were singletons, respectively. New repeat boxes within the *flaA* sequences were found in this study. Unweighted Pair Group Method with Arithmetic Mean dendrogram based on nucleotides of the full length *flaA* gene, the *flaA* short variable region gene and the in silico *flaA* phylogenetic tree of *DdeI* restriction fragment length polymorphism (RFLP) profiles produced very similar clustering with a diversity index of 0.86 for each of the 3 methods. We conclude that *flaA* typing based on *DdeI* RFLP of the PCR products is a cheap, rapid, and reliable method for the epidemiological study of *C. jejuni* isolates of clinical origin in resource-limited regions or in large-scale population surveillance.

INTRODUCTION

The genus *Campylobacter* belongs to the family of *Campylobacteraceae* and includes some of most important intestinal pathogenic bacteria. The bacteria from this genus cause diarrhea and systematic diseases; they are important infectious agents in children less than 5 years of age and in young adults in both developed and developing countries (1,2). Within the genus *Campylobacter jejuni* is the most common pathogenic species in humans and is one of the most important causative agents of acute gastro-enteritis. This bacterium is transmitted to humans via consumption of contaminated foods and contact with warm-blooded animals, especially poultry, making epidemiological studies and source tracking imperative (3,4).

DNA sequencing techniques and the availability of sequence data of pathogenic bacteria via online databases allow researchers to perform comparative genomics studies. The sequence-based methods have high sensitivity, rapidity, discriminatory power, and reproducibility, and provide the interpretation and standardized nomenclature of the subtypes, which permit inter-laboratory comparisons and electronic distribution (5–7). Several methods have been used for *C. jejuni* and *Campylobac-*

ter coli genotyping, including pulsed field gel electrophoresis (PFGE), multilocus sequence typing (MLST), and amplified fragment length polymorphism (AFLP). Flagellin gene restriction fragment length polymorphism (*flaA*-RFLP) has achieved both epidemiological and strain discrimination goals in outbreak analyses (8–10).

flaA sequence typing is a genotyping method that is increasingly used in epidemiological investigations of thermo-tolerant *Campylobacter* species (11,12). However, for *C. jejuni*, RFLP may be just as discriminatory as PFGE and MLST; moreover, it is assuredly more convenient than the latter 2 techniques. The New England Biolab (NEB) cutter online software can be used for computer-simulated RFLP analysis and virtual gel-plotting of *flaA* alleles. To the best of our knowledge, there are no data on the molecular epidemiology of *C. jejuni* based on *flaA* gene sequencing analyses among clinical *C. jejuni* in Iran. This study will provide insight into the genetic relatedness of Iranian clinical *C. jejuni* isolates based on *flaA* sequence. We explored the use of DNA sequencing of the *flaA* gene and short variable region (SVR)-*flaA* together with computer-simulated RFLP analysis based on *flaA* sequence divergence as basis for subtyping of *C. jejuni* isolates of human origin.

MATERIALS AND METHODS

Study design and data collection procedure: Four hundred stool specimens were collected from diarrheal patients attending to 3 major children's hospitals and medical centers in 8 different provinces of Iran between July 2012 and June 2013. Children under 5 years of age with acute diarrhea were included in this study; those

Received August 11, 2015. Accepted December 16, 2015.
J-STAGE Advance Publication January 8, 2016.
DOI: 10.7883/yoken.JJID.2015.386

*Corresponding author: Mailing address: Department of Bacteriology, Faculty of Medical Sciences, Tarbiat Modares University, Jalal Al Ahmad Ave., Tehran 14117-13116, Iran. Tel: +98-2182884558, Fax: +98-2182884555, E-mail: b.bakhshi@modares.ac.ir

with persistent diarrhea or those who had received previous treatment with antibiotics in the previous 5 days were excluded. Acute diarrhea was defined as more than 3 loose episodes/day (13). All samples were transported to the laboratory in modified Cary-Blair transport medium with reduced agar content (1.6 g/L) (14). Identification of *C. jejuni* was performed by standard conventional biochemical tests and confirmed by PCR amplification of *cadF*, *hipO*, and *aspA*, specific for genus *Campylobacter*, *C. jejuni* and *C. coli*, respectively (15).

Amplification and sequencing of the *flaA* gene: In this study, a novel primer pair was designed based on the conserved sequences flanking the *flaA* locus of *C. jejuni* strain (GenBank accession no. AF050186.1), which encompass and amplify the entire *flaA* gene together with the flanking regions. A forward primer *flaA-F* (5'-TTTCGTATTAACACAAATGGTGC-3') and reverse primer *flaA-R* (5'-CTGTAGTAATCTTAA AACATTTTG-3') were used, and the fragment between 2 primers was amplified in a 25 μ L reaction mixture containing 10 ng of DNA template, 2.5 μ L of the PCR buffer 10 \times , 200 μ M of dNTPs, 5 mM of MgCl₂, 0.1 μ M of each primer, and 1 U of Taq DNA polymerase. The PCR products were sequenced (Genfanavar-Macrogen, Seoul, Korea) using ABI 3730XL capillary sequencer (Applied Biosystems, Foster City, CA, USA) after purification with a QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany).

Computer analysis of the sequence: Bidirectional sequencing of amplified fragments was performed and analyzed using MEGA software version 6.0.6 <<http://megasoftware.net>>. Aligned sequences were compared and a dendrogram was generated using the Unweighted Pair Group Method with Arithmetic mean (UPGMA) method. A similarity cut-off value of 90% was used as

the identity level.

GenBank accession numbers: The twenty nucleotide sequences obtained in this study were deposited in the GenBank database. The assigned GenBank accession nos. are KM396335 to KM396354.

In silico restriction analysis pattern: The aligned and trimmed 1,681 bp-fragment sequences were digested in silico with *DdeI* restriction enzyme (restriction enzymes are endonucleases that digest DNA at specific palindromic sites) and exported to the computer-simulated RFLP analysis and virtual gel plotting program (NEB cutter online software version: 1.0.0.4028 <<http://www.neb.com>>). The UPGMA dendrogram was drawn using GelCompar II software (Applied Maths, Sint-Martens-Latem, Belgium) with a similarity cut-off value of 90% as the identity level.

RESULTS

Bacterial strains: Twenty *Campylobacter* spp. were identified from a total of 400 clinical specimens taken from patients with diarrheal symptoms in Iran. PCR-based identification determined that all the *Campylobacter* isolates collected herein comprised *C. jejuni*.

Total *flaA* and *flaA*-SVR nucleotide and amino acid sequence analysis: The designed PCR primers used in this study amplified 1,743 bp of *flaA* gene in all (100%) of our *Campylobacter* spp. isolates. After splicing the low-quality sequences at the ends, a region of 1,681 bp, which harbored the entire *flaA* sequence and produced 560 amino acids was subjected to analysis. The *flaA* sequences of 20 isolates were aligned and compared with that of *C. jejuni* strain NCTC11168, the complete genome of which is available in GenBank (Accession no. AF050186.1). The nucleotide composition of 1,681 bp *flaA* segments showed an average G + C content of

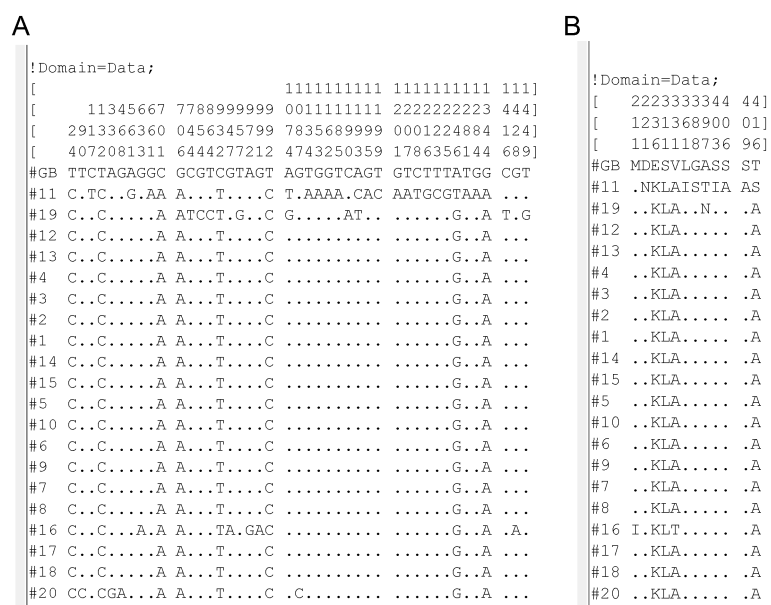


Fig. 1. (A) Forty three nucleotide singleton sites within the 1,681-bp segment of the *flaA* gene among 20 clinical isolates of *C. jejuni*. (B) Twelve amino acid singleton sites among clinical isolates of *C. jejuni*. Numbering of the polymorphic sites is in the vertical format and the name of each isolate is indicated in the left. The numbers in the upper panel are the position of nucleotides (A) and amino acids (B) in *flaA* gene. GB, GenBank reference sequence with which sequences are compared.

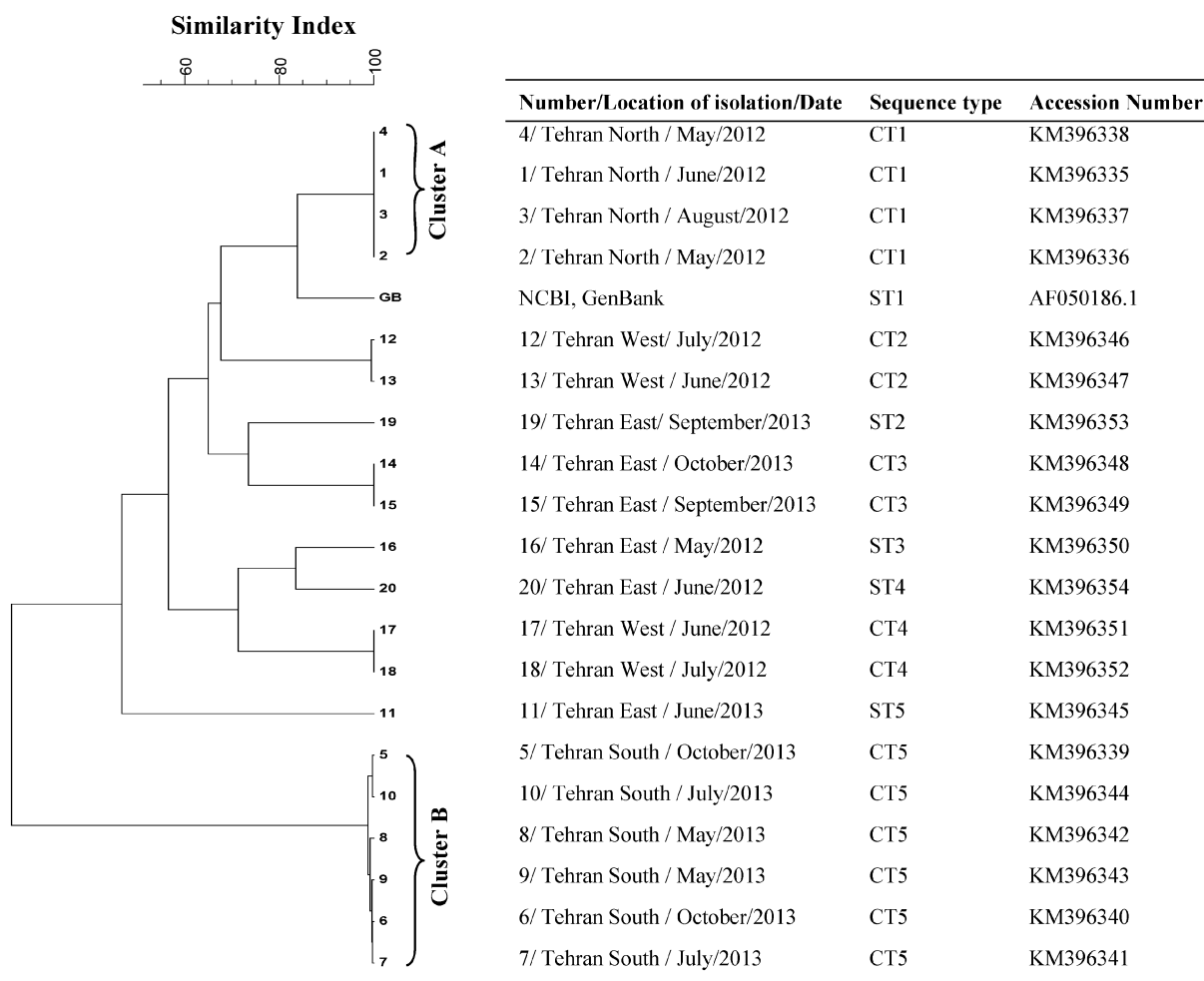


Fig. 2. UPGMA dendrogram made from 20 clinical *C. jejuni* isolates based on the nucleotide similarity of 1,681 nucleotides from total *flaA* sequences. *C. jejuni* strain NCTC11168 (GenBank accession no. AF050186.1) was used as a control. The data are depicted in relation to location, date of isolation, sequence type (CT, common type; ST, single type), and related Accession no. Horizontal axis of dendrogram shows the similarity index among the isolates.

36.8%. A total of 372 (22.1%) nucleotide variable sites and 90 (16.1%) amino acid variable sites were obtained from the isolates under scrutiny. Moreover, 43 (2.5%) and 12 (0.7%) singleton sites were detected in the nucleotide and amino acid sequences, respectively (Fig. 1A and B). No cysteine, histidine, and tryptophan residues were recognized in the *flaA* sequences. Proline residues accounted for only 0.08% of the total amino acids, whereas, serine, glycine, and alanine accounted for 11.99%, 11.43%, and 10.77% of the amino acids, respectively.

The *flaA* amino acid sequence contained 2 perfect repeat boxes at positions 404–451. The 1st box comprised glycine, serine, glycine, phenylalanine, serine, alanine/serine, glycine, and serine (GSGFSA/SGS) and the 2nd box comprised glycine, serine, glycine, phenylalanine, serine, serine, glutamine/glycine, and serine (GSGFSSQ/GS). These boxes existed in the carboxyl end of the *flaA* genes of all of the *C. jejuni* isolates.

Phylogenetic relationships in the population: A UPGMA dendrogram was created based on the sequence homology of the *flaA* gene in the isolates. Two major clusters were obtained (A and B) with 4 and 6 members,

respectively. All members in cluster A were composed of identical sequences, whereas differences in 8 nucleotide positions, resulting in 5 amino acid changes, were present in the isolates in cluster B (Fig. 2). Interestingly, all isolates from Tehran North and Tehran South were included in clusters A and B, respectively, despite discrete isolation dates. The most parsimonious relationships were analyzed; the level of discrimination that could be discerned from the entire sequence of the total *flaA* gene was compared with the *flaA*-SVR and the results were very similar (Fig. 3). In silico *flaA* analysis with the *DdeI* restriction enzyme using NEB cutter online software showed 6–11 fragments. The UPGMA dendrogram from virtual gel plotting among the isolates showed very similar clustering (Fig. 3).

All three methods (*flaA* sequence typing, *flaA*-SVR sequence typing, and PCR-RFLP of the *flaA* sequence) produced identical numbers of profiles (10 profiles), common types (5 CT), and Simpson Diversity indices (Di: 0.86).

DISCUSSION

In the current study, 2 molecular typing methods,

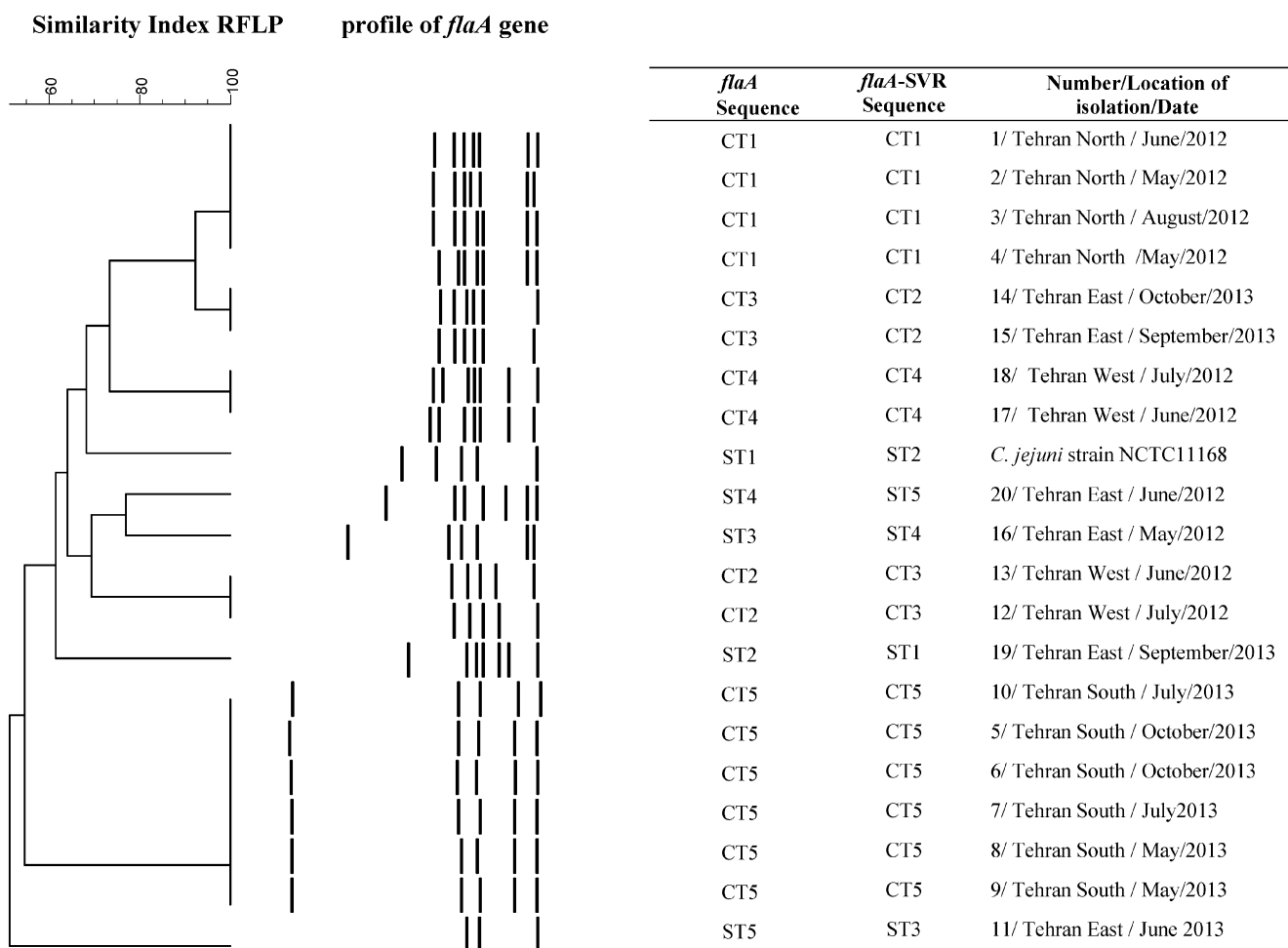


Fig. 3. UPGMA dendrogram made from 20 clinical *C. jejuni* isolates based on in silico PCR-RFLP of *flaA* gene. *C. jejuni* strain NCTC11168 (GenBank accession no. AF050186.1) was used as a control. The data are depicted in relation to location, date of isolation, *flaA*, and *flaA*-SVR sequence types (CT, common type; ST, single type). Horizontal axis of dendrogram shows the similarity index among the isolates.

flaA sequencing and in silico PCR-RFLP, were used to compare relative discriminatory power and typability for differentiation in 20 *C. jejuni* isolates of clinical origin. A *DdeI* restriction enzyme was used for in silico restriction digestion of the complete *flaA* gene; *DdeI* is supposed to provide distinguishable banding patterns in comparison with other restriction enzymes including *Sau3AI* and *AluI* (16,17).

The flagellum is a target for the human immune system, and any mutation event in the *flaA* locus that engenders positive selection increases the growth or survival of the bacterium (18). Comparison of a 1,681-bp trimmed nucleotide sequence of the *flaA* gene from 20 isolates of *C. jejuni* showed 2 highly variable regions; one at positions 539 to 689 designated the *flaA*-SVR; and a large variable region (*flaA*-LVR), at approximately positions 700 to 1,600. The SVR region of the *flaA* gene has been previously described by Khawaja et al. and Meinersmann et al. at base positions 450 to 600 (19,20). This region showed a more narrow range of diversity in the *flaA*-SVR among the *Campylobacter* population under study. The sequence data of the LVR showed similar results to those of the SVR; owing to the length of the sequence and the complexity of LVR data analysis, this region is not suitable for molecular epi-

demological studies of *Campylobacter* infections.

The average G + C content of the *flaA* 1,681-bp segments was 36.8%, which is consistent with the equivalent data reported by Khawaja et al. (36.4% G + C content of the *flaA* gene in *C. jejuni* TGH9011 (19)). This phenomenon shows that, despite continuous point mutations occurring within the coding sequence of the *flaA* gene, the overall nucleotide composition of the gene has not substantially changed among different populations. The *flaA* amino acid sequence contained 2 perfect repeat boxes (GSGFSA/SGS and GSGFSSQ/GS) in 2 positions within the carboxyl end in all of our *C. jejuni* isolates. To the best of our knowledge, although repeat boxes have been introduced previously, boxes with such sequences were introduced for the first time in this study.

A total of 22% nucleotide variable sites versus 16% amino acid variable sites were obtained among the isolates under scrutiny which shows that 6% of the mutations are silent. Moreover 2.5% singleton sites were detected in the nucleotide sequences, all of which were effective in changing the related amino acid sequences. In silico RFLP analysis of the *flaA* gene according to sequence data obtained through total *flaA* sequencing generated a UPGMA dendrogram similar to that gener-

ated following analysis of the complete *flaA* gene sequence. This suggests that PCR-RFLP analysis can be considered an effective genotyping tool in epidemiological investigations where financial resources are limited or in large-scale population surveillance. Parsimony analysis of the *flaA*-SVR sequences also generated a UPGMA dendrogram that was closely related to that generated by analysis of the complete *flaA* gene sequence. To the best of our knowledge, this study is the first report to compare 3 *flaA* typing methods. Accordingly, comparison of *flaA*-SVR sequencing and *flaA*-RFLP typing methods suggests that, although the 2 methods discriminate slightly differently among *Campylobacter* populations, both are effective for outbreak analyses and source-tracking studies (8).

In contrast to the 2 ends of the flagellin gene of *C. jejuni*, the sequence of the related protein within its internal regions was highly conserved. Such patterns were also obvious in the flagellin sequence of *C. coli* and other bacteria such as *Bacillus subtilis*, *Salmonella typhimurium*, and *Escherichia coli*. It can be proposed that the mutational and/or recombination events at the ends of the *flaA* gene, occur more frequently, which accounts for the antigenic variation of this region (21–23).

In conclusion, this study showed; i) analysis of the *flaA*-SVR yielded very similar results to those obtained by sequencing the entire gene; ii) *flaA*-SVR typing is a cheap, rapid, and reliable method for epidemiological study of *C. jejuni* isolates of clinical origin.

Acknowledgments We thank the Research Council of Tarbiat Modares University for supporting the project.

Conflict of interest None to declare.

REFERENCES

- Randall L, Lemma F, Rodgers J, et al. Development and evaluation of internal amplification controls for use in a real-time duplex PCR assay for detection of *Campylobacter coli* and *Campylobacter jejuni*. *J Med Microbiol*. 2010;59:172-8.
- Galanis E. *Campylobacter* and bacterial gastroenteritis. *CMAJ*. 2007;177:570-1.
- Sarkar SR, Hossain MA, Paul SK, et al. *Campylobacteriosis*—an overview. *Mymensingh Med J*. 2014;23:173-80.
- Mohan V. The role of probiotics in the inhibition of *Campylobacter jejuni* colonization and virulence attenuation. *Eur J Clin Microbiol Infect Dis*. 2015;34:1503-13.
- Miller WG. Comparative genomics of *Campylobacter* species other than *Campylobacter jejuni*. In: Nachamkin I, Szymanski C, Blaser CM, editors. *Campylobacter*. 3rd edition. Washington DC: ASM Press; 2008. p. 73-95.
- Sails AD, Swaminathan B, Fields PI. Utility of multilocus sequence typing as an epidemiological tool for investigation of outbreaks of gastroenteritis caused by *Campylobacter jejuni*. *J Clin Microb*. 2003;41:4733-9.
- Cha I, Kim N-O, Nam J-G, et al. Genetic diversity of *Campylobacter jejuni* isolates from Korea and travel-associated cases from east and southeast Asian countries. *Jpn J Infect Dis*. 2014; 67:490-4.
- Clark CG, Bryden L, Cuff WR, et al. Use of the oxford multilocus sequence typing protocol and sequencing of the flagellin short variable region to characterize isolates from a large outbreak of waterborne *Campylobacter* sp. strains in Walkerton, Ontario, Canada. *J Clin Microb*. 2005;43:2080-91.
- Islam Z, van Belkum A, Wagenaar JA, et al. Comparative population structure analysis of *Campylobacter jejuni* from human and poultry origin in Bangladesh. *Eur J Clin Microbiol Infect Dis*. 2014;33:2173-81.
- Djordjevic SP, Unicomb LE, Adamson PJ, et al. Clonal complexes of *Campylobacter jejuni* identified by multilocus sequence typing are reliably predicted by restriction fragment length polymorphism analyses of the *flaA* gene. *J Clin Microb*. 2007;45: 102-8.
- Wassenaar TM, Newell DG. Genotyping of *Campylobacter* spp. *Appl Environ Microbiol*. 2000;66:1-9.
- Kittl S, Heckel G, Korczak BM, et al. Source attribution of human *Campylobacter* isolates by MLST and *fla*-typing and association of genotypes with quinolone resistance. *PLoS One*. 2013; 8:e81796.
- Bhutta ZA, Bird SM, Black RE, et al. Therapeutic effects of oral zinc in acute and persistent diarrhea in children in developing countries: pooled analysis of randomized controlled trials. *Am J Clin Nutr*. 2000;72:1516-22.
- Isenberg HD. *Clinical Microbiology Procedures Handbook*. Washington DC: ASM Press; 1992.
- Ghorbanalizadgan M, Bakhshi B, Lili AK, et al. A molecular survey of *Campylobacter jejuni* and *Campylobacter coli* virulence and diversity. *Iran Biomed J*. 2014;18:158-64.
- Harrington CS, Moran L, Ridley AM, et al. Inter-laboratory evaluation of three flagellin PCR/RFLP methods for typing *Campylobacter jejuni* and *C. coli*: the CAMPYNET experience. *J Appl Microbiol*. 2003;95:1321-33.
- Khoshbakht R, Tabatabaei M, Hosseinzadeh S, et al. Genetic characterization of *Campylobacter jejuni* and *C. coli* isolated from broilers using *flaA* PCR-restriction fragment length polymorphism method in Shiraz, southern Iran. *Jundishapur J Microbiol*. 2015;8:e18573.
- Haiko J, Westerlund-Wikström B. The role of the bacterial flagellum in adhesion and virulence. *Biology (Basel)*. 2013;2: 1242-67.
- Khawaja R, Neote K, Bingham HL, et al. Cloning and sequence analysis of the flagellin gene of *Campylobacter jejuni* TGH9011. *Curr Microbiol*. 1992;24:213-21.
- Meinersmann R, Hessel L, Fields P, et al. Discrimination of *Campylobacter jejuni* isolates by *fla* gene sequencing. *J Clin Microbiol*. 1997;35:2810-4.
- Singh P, Kwon YM. Comparative analysis of *Campylobacter* populations within individual market-age broilers using *fla* gene typing method. *Poult Sci*. 2013;92:2135-44.
- Konkel ME, Christensen JE, Dhillon AS. *Campylobacter jejuni* strains compete for colonization in broiler chicks. *Appl Environ Microbiol*. 2007;73:2297-305.
- Lüneberg E, Glenn-Calvo E, Hartmann M, et al. The central, surface-exposed region of the flagellar hook protein FlgE of *Campylobacter jejuni* shows hypervariability among strains. *J Bacteriol*. 1998;180:3711-4.