Molecular analysis and characterisation of the full-length flagellin C gene (*flaC*) from *Campylobacter lari*

M. MURAYAMA*, T. SEKIZUKA*, A. TAZUMI*, J. E. MOORE†, B. C. MILLAR† and M. MATSUDA*

Laboratory of Molecular Biology, School of Environmental Health Sciences, Azabu University, Fuchinobe, Sagamihara, Japan; and †Department of Bacteriology, Northern Ireland Public Health Laboratory, Belfast City Hospital, Belfast, Northern Ireland, UK

Accepted: 26 October 2010

ABSTRACT

A degenerate polymerase chain reaction (PCR) primer pair (f-Cl*flaC*/r-Cl*flaC*) was constructed *in silico* to amplify *flaC* and its adjacent genetic loci from *Campylobacter lari* isolates. Approximately 1.45 kbp amplicons, including the sequences encoding the *flaC* structural gene of 750 bp, putative promoter, ρ-independent intrinsic terminator regions and partial sequences of two putative open reading frames (ORFs), immediately upstream and downstream of the gene, were identified in 16 *C. lari* isolates (four urease-negative [UN] *C. lari*; 12 urease-positive thermophilic campylobacters [UPTC)]). All 16 *flaC* structural genes commenced with an ATG start codon and terminated with a TAA stop codon and probable ribosome-binding sites were identified in all 16 isolates. These probably indicate a monocistronic operon structure for the *flaC* gene in *C. lari* isolates. In addition, the putative *flaC* gene ORFs were deduced to be similar in 747 bp among all 26 thermophilic *Campylobacter* isolates examined, resulting in a similar calculated molecular weight of approximately 26.6–26.9 kDa. The *flaC* from *C. lari* was different from the *flaA*-like sequence and the shorter *flaA* of UPTC isolates found previously. Reverse transcription PCR and Northern blot hybridisation analyses identified *flaC* transcription in *C. lari* cells. The transcription initiation site for the *flaC* gene was also determined by primer extension analysis. A dendrogram constructed, based on the nucleotide sequence information of *flaC* from 17 *C. lari* isolates, demonstrated that the *C. lari* isolates were genetically variable and formed two minor clusters for UN *C. lari* and UPTC.

KEY WORDS: Campylobacter lari.
Cloning.
Phylogenetic analysis.
Sequencing.

Introduction

Campylobacter lari was first isolated from mammalian and avian species, particularly seagulls of the genus *Larus*.[1,2] *C. lari* has also been shown to be a cause of clinical infection.[3,4]

An atypical group of 10 isolates of urease-positive thermophilic *Campylobacter* (UPTC) was first isolated from the natural environment in England in 1985.[5] Thereafter, these organisms were described as a biovar or variant of *C. lari*,[6,7] and Megraud and his colleagues described four human isolates in France.[6,8] Additional isolates of UPTC have been reported in Northern Ireland,[9-11] in The Netherlands,[12] and in Japan.[13,14] Thus, these two representative taxa, namely urease-negative (UN) *C. lari* and UPTC, occur within the species *C. lari*.[15]

Bacterial flagella are important for the motility and chemotaxis of campylobacters.[16] The flagella have been shown to play an important role in adhesion and have been reported to be one of the best-defined virulence factors in certain enteropathogenic bacteria, including *C. jejuni*.[17]

Recently, Song *et al.*[18] reported the characterisation of a flagellin C gene (*flaC*) encoding a 26 kDa protein (249 amino acid residues, designated FlaC) of the *C. jejuni* TGH9011 strain. FlaC exhibits sequence homology to the N- and C-terminal regions of FlaA and FlaB but lacks the central domain. In addition, it is not required for flagellin morphology or motility. FlaC protein secreted through the flagellar apparatus from *C. jejuni* TGH9011 binds epithelial cells and influences cell invasion.[18] They also suggested that an homologous sequence of *flaC* is present in *C. coli, C. lari* and *C. upsaliensis*, as demonstrated by Southern blot hybridisation analysis, but data were not shown.[18]

The current research group has already demonstrated the *flaA*-like sequences[19] and the shorter *flaA* genes[20] in UPTC isolates. Based on the nucleotide sequence alignment analysis, *flaA*-like sequences and the shorter *flaA* genes appear to be distinctly different from *flaC* in *C. jejuni* TGH9011, described by Song *et al*.[18] However, no analysis of *flaC* from *C. lari* has yet been reported. Therefore, it is important to study *flaC* from *C. lari* isolates (UN *C. lari* and UPTC) and compare it with *flaA*-like sequences and the shorter *flaA* of UPTC isolates and with *flaC* from *C. jejuni*.

Moreover, it would also be worthwhile to construct a polymerase chain reaction (PCR) primer pair *in silico* for amplification of the full-length *flaC* gene, including the putative promoter and terminator regions, and to clone,

Correspondence to: Professor Motoo Matsuda
Laboratory of Molecular Biology, School of Environmental Health Sciences, Azabu University, Fuchinobe 1-17-71, Sagamihara 229-8501, Japan
Email: matsuda@azabu-u.ac.jp

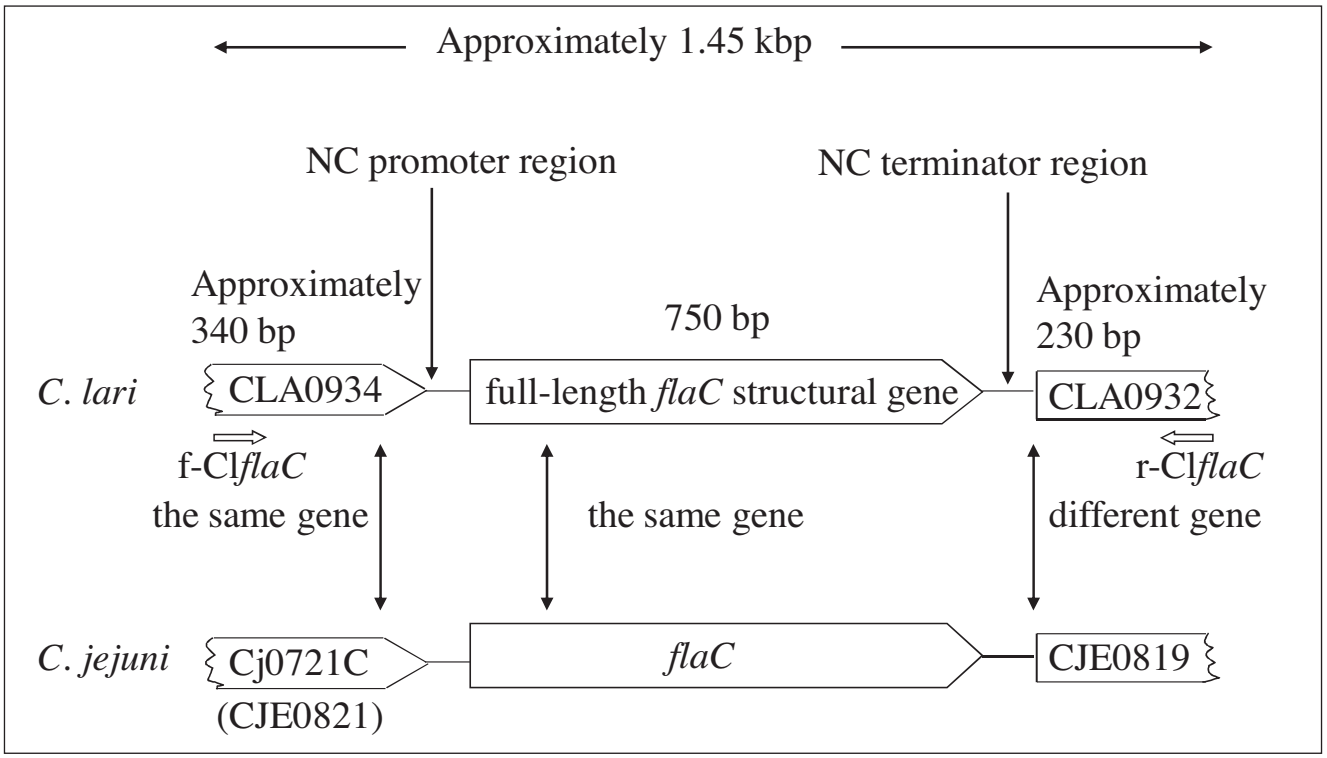

Fig. 1. A schematic representation of the full-length putative *flaC* gene arrangement and its adjacent genetic loci in *C. lari* RM2100 and *C. lari* isolates, including the locations of a primer pair for PCR amplification. The gene arrangement for *C. jejuni* RM1221 strain is also shown for comparison. NC: non-coding.

sequence and analyse the amplicons, as a first step to clarify the *flaC* gene arrangement in *C. lari* organisms.

As FlaC plays an important role in cell invasion,[18] it would be worthwhile to clarify the molecular characteristics of FlaC in *C. lari* organisms, in order to understand the pathogenesis of *C. lari*, which is also shown to be a cause of clinical infection.[3,4]

The present study aims to perform PCR cloning of the full-length *flaC* gene from *C. lari* isolates using a primer pair constructed *in silico* based on the sequence information of the *flaC* genes deposited in the DDBJ/EMBL/GenBank, and to sequence, analyse and characterise this gene.

Materials and methods

Bacterial isolates and culture conditions

C. lari isolates (four UN *C. lari*, 12 UPTC) were used in the present study (Table 1). These isolates were cultured on Mueller-Hinton broth at 37˚C for 48 h in an aerobic jar on Blood Agar Base No.2 (Oxoid, Hampshire, UK) containing 7% (v/v) defibrinated horse blood (Nippon Bio-Test, Tokyo, Japan) and *Campylobacter*-selective medium (Virion, Zurich, Switzerland). An atmosphere of 5% (v/v) O_2 and 10% (v/v) CO_2 was produced by BBL CampyPak Microaerophilic system envelopes (Becton Dickinson, NJ. USA).

Primer design, template DNA preparation and PCR amplification

A degenerate PCR primer pair was designed of f-Cl*flaC* (5′-TTCAAAGYGGVYTWWTRATGA-3′) and r-Cl*flaC* (5′-AAASTTTGCCATYCTSTATCT-3′) *in silico* for amplification of the full-length *flaC* gene, including the putative transcriptional promoter and terminator regions (an approximate 1.45 kbp product) based on the sequence information of the *flaC* gene and the adjacent genetic loci of the gene of *C. lari* RM2100 (AAFK01000002)[21] and other thermophilic campylobacters (Table 1). Nucleotide sequence alignment analysis was performed using the CLUSTAL W software (1.7 program)[22] incorporated in DDBJ.

Template DNA was prepared using sodium dodecyl sulphate and proteinase K treatment, phenol-chloroform extraction and ethanol precipitation.[23] The PCR mixtures contained 50 ng template DNA, 10 mmol/L Tris-HCl (pH 8.3), 50 mmol/L KCl, 1.5 mmol/L $MgCl_2$, 200 µmol each dNTP, 0.64 µmol each primer, and a total of one unit of TaKaRa *Thermus aquaticus* (*Taq*) DNA polymerase (TaKaRa Bio, Shiga, Japan). The PCR was performed in 25 µL reaction volumes at 94˚C for 5 min, then 25 cycles of 94˚C for 30 sec, 50.7˚C for 30 sec, and 72˚C for 60 sec, followed by a final extension of 72˚C for 5 min.

Cloning, nucleotide sequencing and sequence analyses

Amplified PCR products were separated by 1.0% (w/v) agarose gel electrophoresis in 0.5xTBE at 100 V and detected by staining with ethidium bromide. The products were purified using a QIAquick PCR purification kit (Qiagen, Tokyo, Japan) and inserted into the pGEM-T vector using the pGEM-T Easy vector system (Promega. Tokyo, Japan). Sequencing of the cloned *flaC* gene fragment was performed with an Hitachi DNA autosequencer (SQ5500EL) using a Thermo Sequenase premixed cycle sequencing kit (Amersham Pharmacia Biotech, Tokyo, Japan). Sequence analysis of the PCR amplicons was carried out using GENETYX-Windows software (version 9; GENETYX, Tokyo, Japan).

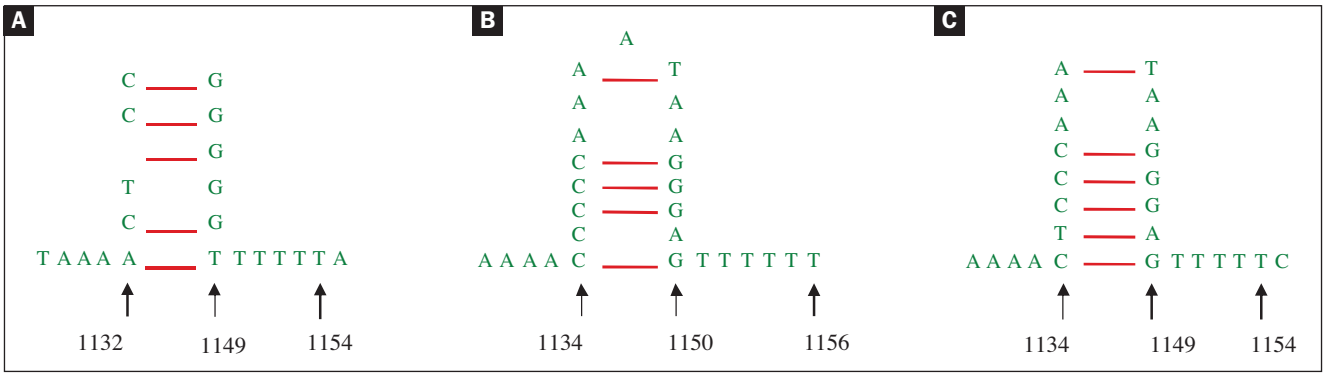

Fig. 2. Three cases of putative intrinsic ρ-independent transcription terminator structures which contains a G+C-rich region near the base of the stem and a single-stranded run of T residues. A) UPTC NCTC12892, B) UPTC 89049, C) *C. lari* JCM 2530^{T}.

Nucleotide sequences of approximately 750 bp of the *flaC* gene from the *C. lari* isolates were compared to each other and to accessible sequence data of other thermophilic campylobacters shown in Table 1, using CLUSTAL W software.[22]

Reverse transcription-PCR analysis and Northern blot hybridisation

Total cellular RNA was extracted and purified from *C. lari* cells using RNA Protect Bacteria Reagent and the RNeasy Mini Kit (Qiagen). Reverse transcription (RT)-PCR analysis was carried out with the primer pair f-/r-Cl*fla*Csg-2 (f-Cl*fla*Csg-2, 5′-GCTGATTCTTTAAGAAGTC-3′, nucleotide positions [np] 513–531 bp; r-Cl*fla*Csg-2, 5′-GCATTTTCTTTTARRTARTTKGC-3′, 1038–1060 bp) using a one-step RT-PCR kit (Qiagen). Northern blot hybridisation analysis was carried out according to the procedure described by Sambrook and Russell,[23] using a PCR- amplified fragment as a probe. The fragment was amplified using a primer pair of f-/r- Cl*fla*Csg-1 (f-Cl*fla*Csg-1, 5′-

Table 1. Isolates of *C. lari* used in the present study, other reference strains, accession numbers of the nucleotide sequence data of the full-length *flaC* gene accessible in the DDBJ/EMBL/GenBank and the putative ORF and CMW of the gene.

Isolate	Source	Country	Accession number	ORF (bp)	Number of AA	CMW (Da)
UN *C. lari* JCM2530^{T}	Seagull	Japan	AB443583	747	249	26,897
UN *C. lari* 298	Human	Japan	AB443584	747	249	26,897
UN *C. lari* 300	Seagull	USA	AB443585	747	249	26,897
UN *C. lari* 84C-1	Human	Northern Ireland	AB443586	747	249	26,867
UPTC 99	Sea water	Northern Ireland	AB443587	747	249	26,813
UPTC NCTC12892	River water	England	AB266773	747	249	26,829
UPTC NCTC12893	River water	England	AB266777	747	249	26,829
UPTC NCTC12894	Sea water	England	AB266774	747	249	26,813
UPTC NCTC12895	Mussel	England	AB266775	747	249	26,795
UPTC NCTC12896	Mussel	England	AB266776	747	249	26,813
UPTC CF89-12	River water	Japan	AB266772	747	249	26,814
UPTC A1	Seagull	Northern Ireland	AB266769	747	249	26,857
UPTC A2	Seagull	Northern Ireland	AB266770	747	249	26,857
UPTC A3	Seagull	Northern Ireland	AB266771	747	249	26,827
UPTC 89049	Human	France	AB266767	747	249	26,827
UPTC 92251	Human	France	AB266768	747	249	26,827
C. lari RM2100	Human	USA	AAFK01000002	747	249	26,867
C. jejuni NCTC11168	Human	UK	NC_002163	747	249	26,606
C. jejuni RM1221	Chicken	USA	NC_003912	747	249	26,615
C. jejuni 81-176	Human	USA	NC_008787	747	249	26,574
C. jejuni 260.94	Human	South Africa	AANK01000004	747	249	26,575
C. jejuni CF93-6	Human	Japan	AANJ01000005	747	249	26,624
C. jejuni HB93-13	Human	China	AANQ01000001	747	249	26,569
C. jejuni 84-25	Human	Unknown	AANT02000001	747	249	26,606
C. coli RM2228	Chicken	USA	AAFL01000008	747	249	26,669
C. upsaliensis RM3195	Human	USA	AAFJ01000005	747	249	26,647

Table 2. Nucleotide (upper right) and deduced amino acid (lower left) sequence similarities (%) of the putative ORFs of *flaC* among *C. lari*, *C. jejuni*, *C. coli* and *C. upsaliensis* isolates.

	1	2	3	4	5	6	7	8	9	10	11
1ÅA UN *C. lari* JCM2530^T		99.2	99.2	98.3	91.2	90.0	90.1	91.3	91.2	91.0	91.8
2ÅA UN *C. lari*298	100.0		100.0	98.5	91.2	89.7	89.8	91.3	91.2	91.0	92.1
3ÅA UN *C. lari* 300	100.0	100.0		98.5	91.2	89.7	89.8	91.3	91.2	91.0	92.1
4ÅA UN CC. *lari* 84C-1	99.6	99.6	99.6		91.6	89.8	90.0	91.7	91.6	91.7	92.5
5ÅA UPTC 99	95.2	95.2	95.2	95.5		93.8	94.0	99.7	99.6	99.7	95.2
6ÅA UPTC NCTC12892	94.4	94.4	94.4	94.0	97.6		99.6	94.1	94.0	93.8	92.1
7ÅA UPTC NCTC12893	94.4	94.4	94.4	94.0	97.6	100.0		94.2	94.1	94.0	92.2
8ÅA UPTC NCTC12894	95.2	95.2	95.2	95.6	100.0	97.6	97.6		99.9	99.7	95.3
9ÅA UPTC NCTC12895	94.8	94.8	94.8	95.2	99.6	97.2	97.2	99.6		99.6	95.2
10ÅA UPTC NCTC12896	95.2	95.2	95.2	95.6	100.0	97.6	97.6	100.0	99.6		95.1
11ÅA UPTC CF89-12	95.2	95.2	95.2	95.6	98.8	96.8	96.8	98.8	98.4	98.8	
12ÅA UPTC A1	94.0	94.0	94.0	94.0	95.2	94.8	94.8	95.2	94.8	95.2	95.2
13ÅA UPTC A2	94.0	94.0	94.0	94.0	95.2	94.8	94.8	95.2	94.8	95.2	95.2
14ÅA UPTC A3	94.4	94.4	94.4	94.4	95.6	95.2	95.2	95.6	95.2	95.6	95.6
15ÅA UPTC 89049	94.4	94.4	94.4	94.4	95.6	95.2	95.2	95.6	95.2	95.6	95.6
16ÅA UPTC 92251	94.4	94.4	94.4	94.4	95.6	95.2	95.2	95.6	95.2	95.6	95.6
17ÅA *C. lari* RM2100	99.6	99.6	99.6	100.0	95.6	94.0	94.0	95.6	95.2	95.6	95.6
18ÅA *C. jejuni* NCTC11168	68.3	68.3	68.3	67.9	65.9	66.3	66.3	65.9	65.5	65.9	65.9
19ÅA *C. jejuni* RM1221	68.7	68.7	68.7	68.3	66.3	66.7	66.7	66.3	65.9	66.3	66.3
20ÅA *C. jejuni* 81-176	68.3	68.3	68.3	67.9	65.9	66.3	66.3	65.9	65.5	65.9	65.9
21ÅA *C. jejuni* 260.94	68.7	68.7	68.7	68.3	66.3	66.7	66.7	66.3	65.9	66.3	66.3
22ÅA *C. jejuni* CF93-6	68.7	68.7	68.7	68.3	66.3	66.7	66.7	66.3	65.9	66.3	66.3
23ÅA *C. jejuni* HB93-13	69.1	69.1	69.1	68.7	66.7	67.1	67.1	66.7	66.3	66.7	66.7
24ÅA *C. jejuni* 84-25	68.3	68.3	68.3	67.9 6	5.9	66.3	66.3	65.9	65.5	65.9	65.9
25ÅA *C. coli* RM2228	68.7	68.7	68.7	68.3	66.3	66.7	66.7	66.3	65.9	66.3	66.3
26ÅA *C. upsaliensis* RM3195	67.5	67.5	67.5	67.5	67.9	68.3	68.3	67.9	67.5	67.9	67.9

GGCGATATAGGAACAACCCC-3′, np 387–405 bp; r-Cl*flaC*sg-1, 5′-GGGCACTTGCAAATAAATACGC-3′, np 1059–1080 bp). In the present study, the nucleotide positions used are for those of the *C. lari* JCM2530T (DDBJ/EMBL/GenBank Accession No. AB443583). Random primer extension was performed in order to prepare the probe using DIG-High Prime (Roche Applied Science, Penzberg, Germany).

Primer extension analysis

The transcription initiation site of the *flaC* gene was determined by primer extension analysis from the purified total cellular RNA of *C. lari* JCM2530^T cells. The primer selected for this analysis was 5′-GACTTCTTAAAGAATCAGC-3′, which corresponds to the reverse complementary sequence of np 513–531 bp. The transcription initiation site was determined by primer extension, with the sizes of DNA fragments generated by sequencing reactions.

Phylogenetic analysis

A phylogenetic tree was constructed by the neighbour-joining (NJ) method.[24]

Accession numbers

Nucleotide and amino acid sequences cloned and sequenced in the present study are accessible in the DDBJ/EMBL/GenBank under the accession numbers shown in Table 1.

Results

A degenerate PCR primer pair (f-/-rCl*flaC*) constructed *in silico* in the present study amplified PCR products of approximately 1.45 kbp in length (data not shown). Then, as a first step to clarify *flaC* gene arrangement in *C. lari* organisms, the sequences of the *flaC* gene of 750 bp, non-coding (NC) putative transcription promoter and ρ-independent terminator regions and two partial and putative open reading frames (ORFs) immediately upstream and downstream of the *flaC* operon (Fig. 1) were identified in all 16 *C. lari* isolates examined. In Figure 1, the known *flaC* gene arrangement from the *C. jejuni* RM1221 strain is shown for comparison. The present nucleotide sequence analysis of *C. lari* JCM2530^T indicates a putative ORF, 747 nucleotides in length (np 378–1124 bp), encoding a protein of 249 amino acid residues with calculated molecular weight (CMW) of approximately 26.9 kDa, with sequence similarity of approximately 69% to the FlaC protein of *C. jejuni*.[18] In addition, the putative ORFs of the *flaC* gene from all 17 *C. lari* isolates including *C. lari* RM2100, as well as those of seven *C. jejuni*, one *C. coli* and one *C. upsaliensis* strains, were deduced to be similar (i.e., 747 bp in length [249 amino acid residues]), resulting in similar CMWs of the putative ORFs to be approximately 26.6–26.9 kDa for all 26 thermophilic *Campylobacter* isolates (Table 1). These CMW values were consistent with that of the 26 kDa protein

12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
90.5	90.5	90.2	90.2	90.6	98.3	70.3	70.5	70.1	70.4	70.4	70.7	70.3	70.6	68.4
90.5	90.5	90.4	90.4	90.6	98.5	70.4	70.7	70.0	70.3	70.5	70.5	70.5	70.8	68.3
90.5	90.5	90.4	90.4	90.6	98.5	70.4	70.7	70.0	70.3	70.5	70.5	70.4	70.8	68.3
91.3	91.3	91.2	91.2	91.4	100.0	69.6	69.9	69.7	69.7	69.7	70.0	69.6	70.5	68.5
91.2	91.2	91.2	91.2	91.3	91.6	70.3	70.6	70.0	70.0	70.4	70.3	70.3	70.5	68.9
90.8	90.8	90.5	90.5	90.9	89.8	70.7	70.9	70.0	70.0	70.8	70.7	70.7	69.4	67.3
90.9	90.9	90.6	90.6	91.0	90.0	70.8	71.1	70.5	70.5	70.9	70.8	70.8	69.6	67.5
91.2	91.2	91.2	91.2	91.3	91.7	70.6	70.8	70.3	70.3	70.7	70.6	70.6	70.4	68.9
91.0	91.0	91.0	91.0	91.2	91.6	70.4	70.7	70.2	70.2	70.6	70.4	70.4	70.3	68.8
91.2	91.2	91.2	91.2	91.3	91.7	70.4	70.7	70.2	70.2	70.6	70.4	70.4	70.5	68.8
91.4	91.4	91.4	91.4	91.6	92.5	69.6	69.9	69.6	69.6	70.2	69.9	69.6	69.7	68.9
	100.0	99.2	99.2	99.9	91.3	69.8	70.1	69.6	69.6	70.0	69.8	69.8	70.9	68.5
100.0		99.2	99.2	99.9	91.3	69.8	70.1	69.6	69.6	70.0	69.8	69.8	70.9	68.5
99.6	99.6		100.0	99.3	91.2	69.6	69.8	69.3	69.3	70.0	69.6	69.6	70.5	68.1
99.6	99.6	100.0		99.3	91.2	69.6	69.8	69.3	69.3	70.0	69.6	69.6	70.5	68.1
99.6	99.6	100.0	100.0		91.4	70.0	70.2	69.7	69.7	70.1	70.0	70.0	71.0	68.7
94.0	94.0	94.4	94.4	94.4		69.6	69.9	69.7	69.7	69.7	70.0	69.6	70.5	68.5
66.3	66.3	66.7	66.7	66.7	67.9		99.6	99.1	99.3	99.6	99.1	100.0	89.7	76.7
66.7	66.7	67.1	67.1	67.1	68.3	99.2		98.9	99.2	99.7	99.2	99.6	89.8	77.0
66.3	66.3	66.7	66.7	66.7	67.9	98.8	98.8		99.7	98.9	99.5	99.1	89.7	77.0
66.7	66.7	67.1	67.1	67.1	68.3	99.2	99.2	99.6		99.2	99.7	99.3	89.8	77.1
66.7	66.7	67.1	67.1	67.1	68.3	98.8	99.6	98.4	98.8		99.2	99.6	89.6	76.9
67.1	67.1	67.5	67.5	67.5	68.7	98.4	99.2	98.8	99.2	98.8		99.1	89.6	77.1
66.3	66.3	66.7	66.7	66.7	67.9	100.0	99.2	98.8	99.2	98.8	98.4		89.7	76.7
66.7	66.7	67.1	67.1	67.1	68.3	94.8	94.8	94.8	95.2	94.4	94.4	94.8		77.1
67.1	67.1	67.5	67.5	67.5	67.5	81.9	82.3	81.5	81.9	82.3	81.5	81.9	81.9	

product of the *flaC* gene in *C. jejuni* TGH9011, as described previously by Song *et al*.[18]

With regard to the *flaC* gene in the *C. lari* isolates, all 17 putative ORFs commenced with an ATG start codon (np 378–380 bp) and terminated with a TAA stop codon (np 1125–1127 bp). Probable ribosome-binding (RB) sites, AGGAG (np 367–371 bp; Shine-Dalgarno [SD] sequence)[25] and a putative intrinsic ρ-independent transcription terminator region were identified in all 17 *C. lari* isolates. Three cases of the putative intrinsic ρ-independent transcription terminator structure which contains a G+C-rich region near the base of the stem and a single-stranded run of T residues are demonstrated in Figure 2 (A: UPTC NCTC12892; B: UPTC 89049; C: *C. lari* JCM2530^T). The start codon of the *flaC* gene is preceded by the SD sequences.

A promoter consensus sequence at the –10 region (TATTTT, Pribnow box; np 313–318 bp) was identified immediately upstream of the putative ORF of *flaC* in all 17 *C. lari* isolates (Fig. 3). In addition, a consensus sequence at the –35 region (TTATCA; np 289–294 bp) was also identified. Thus, *flaC* may be in a monocistronic operon in *C. lari*, as suggested in *C. jejuni* by Song *et al*.[18]

An approximately 340 bp region (np 1–341 bp) upstream of the NC putative promoter consensus sequence for *flaC* in the 16 *C. lari* isolates, sequenced in the present study, was shown to have a nucleotide sequence similarity to the CLA0934 (probable integral membrane protein cj0721C; AAFK01000002), and an approximately 230 bp region (np 1165–1397 bp) immediately downstream of the putative ρ-independent transcription terminator for *flaC* was shown to have sequence similarity to CLA0932 (conserved hypothetical protein cj0976; AAFK01000002). In relation to CLA0934, each ORF of *C. lari* appears to terminate with a TGA stop codon (np 342–344 bp), except for the ORFs of UN *C. lari* 84C-1 and UPTC 92251, which terminate with a TAA (np 355–357 bp for UN *C. lari* 84C-1 and np 348-350 bp for UPTC 92251) (Fig. 3).

Thus, an overlap of approximately 70 bp may occur between the 3′ end region of CLA 0934 and its NC putative terminator region, and the NC promoter region, including a probable RB site for the *flaC* gene and the NC promoter region, and a probable RB site for the *flaC* gene (from np 301 to 370 approximately in Figure 3).

When RT-PCR analysis was carried out for the *flaC* transcript with a primer pair of f-/r- Cl *flaC*sg-2, positive signals were recognised with two isolates of UN *C. lari* JCM2530^T and UPTC CF89-12 (approximately 550 bp; Fig. 4A) examined. Northern blot hybridisation analysis also detected the *flaC* gene transcription in two *C. lari* cells using a PCR-amplified *flaC* fragment (approximately 800 bp) as a probe (Fig. 4B). Thus, *flaC* gene transcription was confirmed in *C. lari* organisms.

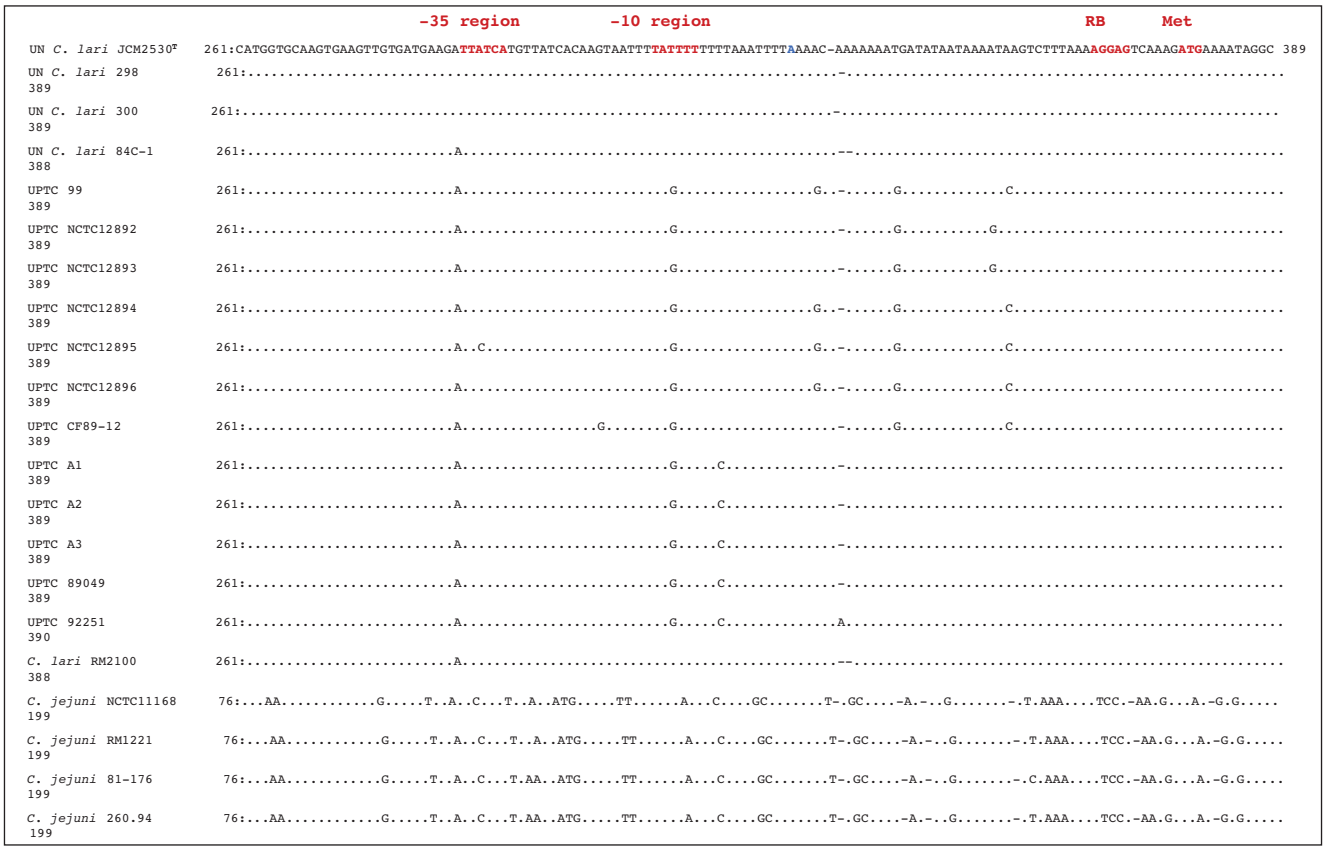

Fig. 3. Nucleotide sequence alignment analyses of the putative promoter structures, consisting of a –35 and a –10 region as well as the start codon ATG. Promoter structures, ribosome binding (RB) sites and start codon in the UN *C. lari* JCM2530^T are in red.

The transcription initiation site for the *flaC* gene was determined by primer extension analysis (Fig. 5). The +1 nucleotide (A) preceding the ATG start codon for the *flaC* gene is underlined in the following sequence: 5′-TTTTAAATTTTAAAACAAA-3′ (Fig. 5).

A phylogenetic tree constructed using the NJ method based on the nucleotide sequence similarity data of the *flaC* structural gene from 16 *C. lari* isolates and other thermophilic campylobacters is shown in Figure 6. There are some major clusters in the dendrogram. The phylogenetic tree also demonstrated that the 16 *C. lari* isolates and a reference strain of UN *C. lari* RM2100 formed minor clusters consisting of one UPTC and one UN *C. lari*, showing genetic hypervariability and separate from other thermophilic campylobacters. Thus, two representative taxa of UN *C. lari* and UPTC were shown to be different, based on the nucleotide sequence information of the *flaC* structural gene.

Discussion

With regard to the *flaC* gene arrangement and its adjacent genetic loci identified in the present study, the same two genes (i.e., CLA0934 and Cj0721C) occurred upstream of the *flaC* gene in 16 *C. lari* isolates examined, as well as in *C. jejuni* (Fig. 1). However, two other genes (i.e., CLA0932 and CJE0819) were different downstream of the *flaC* between these two taxa.

The nucleotide sequences of the putative ORFs of *flaC* from the 17 *C. lari* isolates, including *C. lari* RM2100, showed 89.7–100% sequence similarity (Table 2). The 17 *C. lari* isolates showed 69.3–71.1% similarity to the seven *C. jejuni* strains (Table 2). In addition, the putative ORFs of *flaC* from the 17 *C. lari* isolates also showed 94.0–100% amino acid sequence similarity (Table 2). The 17 *C. lari* isolates showed 65.5–69.1% amino acid sequence similarity to those of the seven *C. jejuni* strains and 65.9–68.7% to those of two *C. coli* and *C. upsaliensis* strains (Table 2).

The authors have already demonstrated the *flaA*-like sequences, incomplete genes or pseudogenes, containing two internal termination codons (TAG) in two Japanese UPTC isolates (CF89-12 and -14) obtained from river water.[19] Moreover, *flaA* genes of UPTC isolates obtained from the natural environment (i.e., river water, sea water and shellfish) (approximately 1.45 kbp in length) and in UN *C. lari* isolates from humans, seagulls, food animals and a mussel (approximately 1.7 kbp) have been demonstrated.[20,26] In addition, the *flaA*-like sequences[19] and the shorter *flaA* genes[20] from UPTC isolates were shown to be markedly shorter, mainly corresponding to approximately np 1140–1400 bp of the large variable region of the structural gene in *flaA* from *C. jejuni* 81116 (J05635).[16] However, the present nucleotide and deduced amino acid sequence alignment analysis indicated that the *flaC* from 16 *C. lari* isolates (four UN *C. lari*, 12 UPTC) was quite different from *flaA*-like sequences[19] and the shorter *flaA*[20] in UPTC isolates (data not shown).

Although Song *et al.* recently described that a homologous sequence of *flaC* is present in *C. lari*, as demonstrated by Southern blot hybridisation analysis, other data were not

shown.[18] Therefore, the present study is the first demonstration of the full-length *flaC* gene, including the occurrence of putative promoter and terminator regions in 16 *C. lari* isolates.

It has been suggested that the secreted form of FlaC is subjected to *N*-linked glycosylation in *C. jejuni*.[18,27] At present, the deduced amino acid sequence of the *flaC* ORF from the 16 *C. lari* (four UN *C. lari*, 12 UPTC) isolates and UN *C. lari* RM2100 strain contained four *N*-glycosylation motifs (Asn-Xaa-Ser/Thr-Xaa; amino acid residues no. 79–82, 124–127, 156–159 and 191–194 for *C. lari* JCM2530^T). Only those of two UPTC isolates, A1 and A2, contained five *N*-glycosylation motifs (i.e., the four motifs described above and one other motif [amino acid residues no. 16–19 for A1]). In addition, those of the seven *C. jejuni* isolates listed in Table 2 clearly contained five *N*-glycosylation motifs (data not shown). Thus, it might be suggested that the number of *N*-glycosylation motifs of the *flaC* may be variable in *C. lari*, although the three sites of the motifs (amino acid residues no. 79–82, 124–127 and 191–194 for NCTC12892) were identical among the four thermophilic *Campylobacter* organisms listed in Table 2 (data not shown).

Although four clinical isolates of UPTC were obtained from humans in France in the period from 1986 to 1989,[6,8] the possible association of UPTC with human and animal disease remains unclear. Therefore, the degenerate PCR primer pair constructed in this study for the specific amplification and detection of the full-length *flaC* gene from *C. lari* isolates would be valuable for the diagnosis of UPTC and UN *C. lari* for medical, veterinary and environmental bacteriology. Thus, additional new sequence information of the full-length *flaC* gene from UN *C. lari* and UPTC isolates would be helpful for *C. lari* diagnosis.

Werno *et al*. identified three thermophilic *Campylobacter* species isolates, *C. jejuni, C. coli* and *C. lari*, by molecular confirmation methods based on a multiplex PCR targeting the *Campylobacter lpxA* gene encoding UDP-*N*-acetyl glucosamine acyltransferase, in a reported on prosthetic devices. As the degenerate primer pair of f-/r-Cl*flaC* in the present study was designed based on the sequence information of the *flaC* gene and its adjacent genetic loci of *C. lari* RM2100 and other thermophilic campylobacters, it might be of value for the differentiation of the thermophilic *Campylobacter* organisms. □

This research was supported partly by the Promotion and Mutual Aid Corporation for Private Schools of Japan, Grant-in-Aid for Matching Fund Subsidy for Private Universities and by a Grant-in-Aid for Scientific Research (C) (No. 20580346) from the Ministry of Education, Culture, Sports, Science and Technology of Japan (to MM). This study was also supported partly by a project grant (Start-Up Support for the Matching Fund Subsidy for Private Universities, 2007–2008) awarded by the Azabu University Research Services Division.

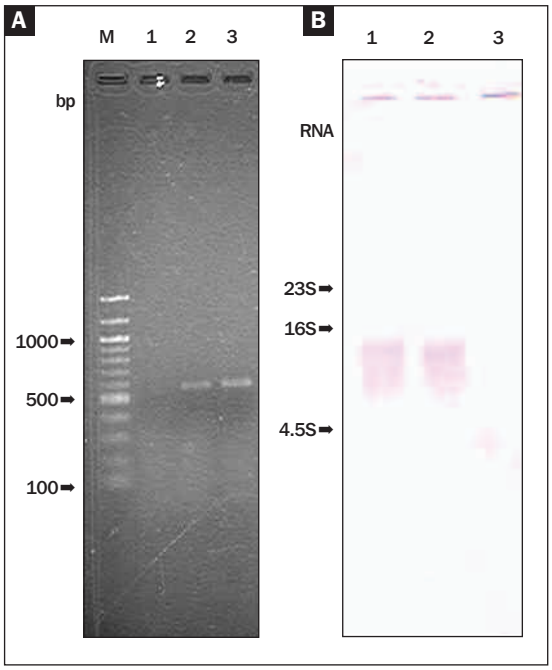

Fig. 4. RT-PCR and Northern blot hybridisation analyses of the *flaC* gene transcript expressed in the *C. lari* isolates. The total RNA components were subjected to electrophoresis. RT-PCR procedures: A) Lane M: 100 bp DNA ladder; lane 1: RT-PCR reaction-negative; lane 2: UN *C. lari* JCM2530^T; lane 3: UPTCCF89-12. Northern blotting: B) Lane 1: UN *C. lari* JCM2530^T; lane 2: CF89-12: lane 3: *E. coli* DH 5α.

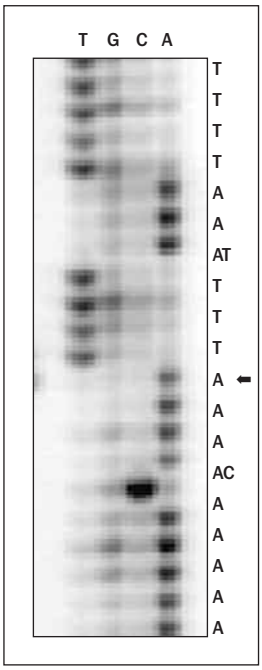

Fig. 5. Primer extension analysis of the *flaC* mRNA transcript in *C. lari* JCM2530^T. The arrow indicates the transcription initiation site.

References

1 Skirrow MB, Benjamin J. '1001' Campylobacters: cultural characteristics of intestinal campylobacters from man and animals. *J Hyg. (Camb)* 1980; **85**: 427–42.

2 Benjamin J, Leaper S, Owen RJ, Skirrow MB. Description of *Campylobacter laridis*, a new species comprising the nalidixic acid-resistant thermophilic *Campylobacter* (NARTC) group. *Curr Microbiol* 1983; **8**: 231–8.

3 Nachamkin I, Stowell C, Skalina D *et al*. *Campylobacter laridis* causing bacteremia in an immunosuppressed patient. *Ann Int Med* 1984; **101**: 55–7.

4 Werno AM, Klena JD, Shaw GM, Murdoch DR. Fatal case of *Campylobacter lari* prosthetic joint infection and bacteremia in an immunocompetent patient. *J Clin Microbiol* 2002; **40**: 1053–5.

5 Bolton FJ, Holt AV, Hutchinson DN. Urease-positive thermophilic campylobacters. *Lancet* 1985; **i**: 1217–8.

6 Mégraud F, Chevrier D, Desplaces N *et al*. Urease-positive thermophilic *Campylobacter* (*Campylobacter laridis* variant) isolated from an appendix and from human feces. *J Clin Microbiol* 1998; **26**: 1050–1.

7 Owen RJ, Costas M, Sloss L, Bolton FJ. Numerical analysis of electrophoretic protein patterns of *Campylobacter laridis* and allied thermophilic campylobacters from the natural environment. *J Appl Bacteriol* 1988; **65**: 69–78.

8 Bezian MC, Ribou G, Barberis-Giletti C, Megraud F. Isolation of a urease-positive thermophilic variant of *Campylobacter lari* from a patient with urinary tract infection. *Eur J Clin Microbiol Infect Dis* 1990; **9**: 895–7.

9 Wilson IG, Moore JE. Presence of *Salmonella* spp. and *Campylobacter* spp. in shellfish. *Epidemiol Infect* 1996; **116**: 147–53.

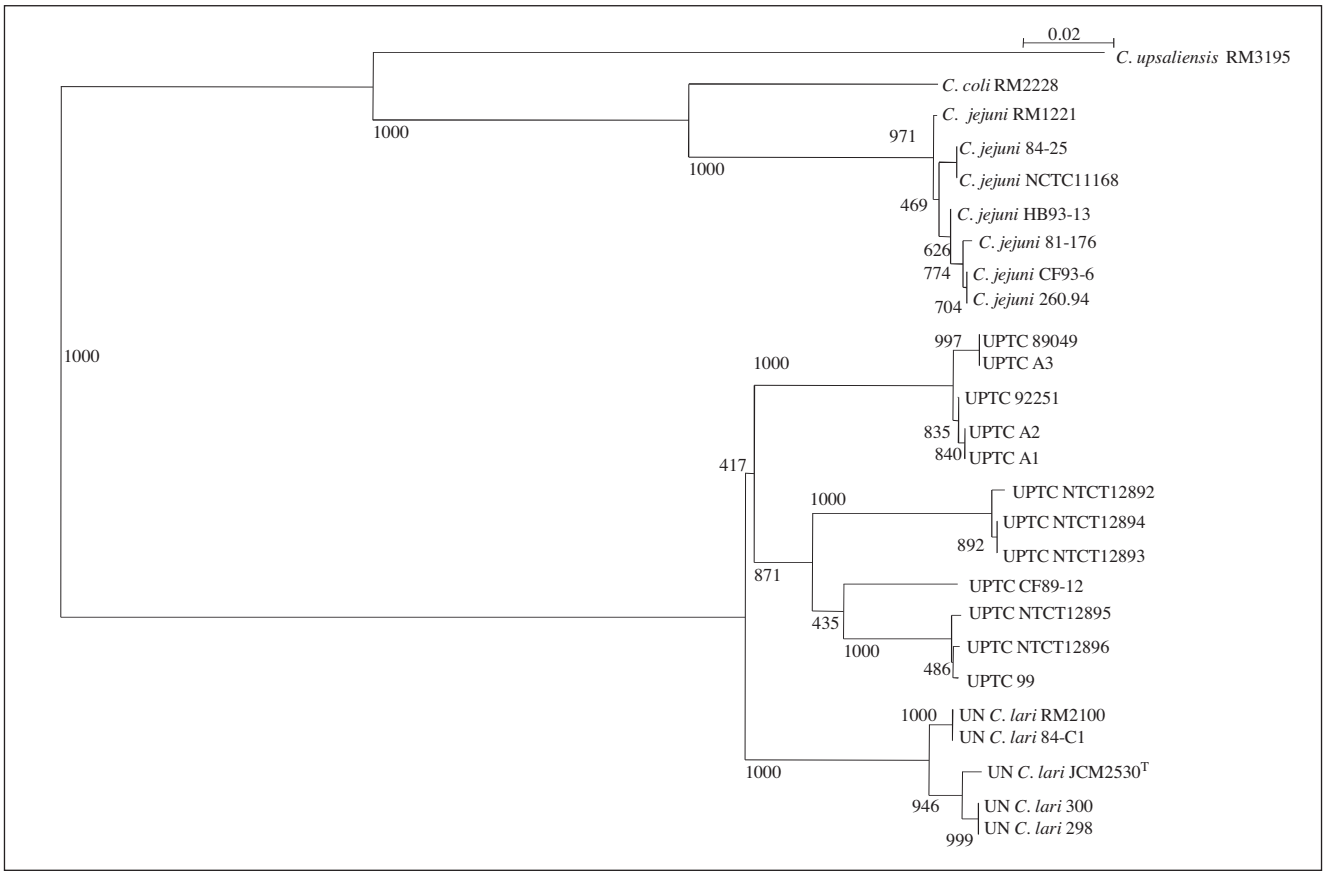

Fig. 6. A phylogenetic tree based on nucleotide sequence information of the *flaC* gene from 17 *C. lari* isolates and other thermophilic campylobacters. The tree was constructed by the NJ method. Values: 0.02 in the figure represents evolutionary distances. Out-group: *C. upsaliensis* RM3195. Boot-strap values of 1000 are shown at the branch points.

10 Kaneko A, Matsuda M, Miyajima M *et al*. Urease-positive thermophilic strains of *Campylobacter* isolated from seagulls (*larus* spp.). *Lett Appl Microbiol* 1999; **29**: 7–9.

11 Matsuda M, Kaneko A, Stanley T *et al*. Characterization of urease-positive thermophilic *Campylobacter* subspecies by multilocus enzyme electrophoresis typing. *Appl Environ Microbiol* 2003; **69**: 3308–10.

12 Endtz HP, Vliegenthart JS, Vandamme P *et al*. Genotypic diversity of *Campylobacter lari* isolated from mussels and oysters in The Netherlands. *Int J Food Microbiol* 1997; **34**: 79–88.

13 Matsuda M, Kaneko A, Fukuyama M *et al.* First finding of urease-positive thermophilic strains of *Campylobacter* in river water in the Far East, namely, in Japan, and their phenotypic and genotypic characterization. *J Appl Bacteriol* 1996; **81**: 608–12.

14 Matsuda M, Shibuya T, Itoh Y *et al.* First isolation of urease-positive thermophilic *Campylobacter* (UPTC) from crows (*Corvus levaillantii)* in Japan. *Int J Hyg Environ Health* 2002; **205**: 321–4.

15 Matsuda M, Moore JE. Urease-positive thermophilic *Campylobacter* species. *Appl Environ Microbiol* 2004; **70**: 4415–8.

16 Nuijten PJ, van Asten FJ, Gaastra W, van der Zeijst BA. Structural and functional analysis of two *Campylobacter jejuni* flagellin genes. *J Biol Chem* 1990; **265**: 17798–804.

17 Logan SM, Harris LA, Trust TJ. Isolation and characterization of *Campylobacter* flagellins. *J Bacteriol* 1987; **169**: 5072–7.

18 Song YC, Jin S, Louie H *et al.* FlaC, a protein of *Campylobacter jejuni* TGH9011 (ATCC43431) secreted through the flagellar apparatus, binds epithelial cells and influences cell invasion. *Mol Microbiol* 2004; **53**: 541–53.

19 Sekizuka T, Gondo T, Murayama O *et al. flaA*-like sequences containing internal termination codons (TAG) in urease-positive thermophilic *Campylobacter* isolated in Japan. *Lett Appl Microbiol* 2002; **35**: 185–9.

20 Sekizuka T, Gondo T, Murayama O *et al*. Molecular cloning, nucleotide sequencing and characterization of the flagellin gene from isolates of urease-positive thermophilic *Campylobacter. Res Microbiol* 2000; **155**: 185–91.

21 Fouts DE, Mongodin EF, Mandrell RE *et al*. Major structural differences and novel potential virulence mechanisms from the genomes of multiple *Campylobacter* species. *PLoS Biol* 2005; **3**: e15 (Epub 4 Jan 2005).

22 Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994; **22**: 4673–80.

23 Sambrook J, Russell DW. *Molecular cloning; a laboratory manual* 3rd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2001.

24 Saitou U, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 1987; **4**: 406–25.

25 Benjamin L. *Genes VII*. Oxford: Oxford University Press, 2000.

26 Sekizuka T, Yokoi T, Murayama O *et al.* A newly constructed primer pair for the PCR amplification, cloning and sequencing of the flagellin (*flaA*) gene from isolates of urease-negative *Campylobacter lari. Antonie Van Leeuwenhoek* 2005; **88**: 113–20.

27 Young NM, Brisson JR, Kelly J *et al.* Structure of the N-linked glycan present on multiple glycoproteins in the Gram-negative bacterium *Campylobacter jejuni. J Biol Chem* 2002; **277**: 42530–9.