Campylobacter lari: molecular and comparative analyses of the virulence-associated chromosome locus J (*vacJ*) gene homologue, including the promoter region

C. TAKAKU*, 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 1-17-71, Sagamihara 229-8501, Japan; and †Department of Bacteriology, Northern Ireland Public Health Laboratory, Belfast City Hospital, Belfast BT9 7AD, Northern Ireland, UK

Accepted: 27 January 2009

ABSTRACT

Following TA cloning and sequencing with a novel *in silico*-designed polymerase chain reaction (PCR) primer pair (f-Cl*vacJ*/r-Cl*vacJ*), approximately 750 base pairs (bp) of promoter and structural gene regions for *vacJ* and its adjacent genetic loci (approximately 1.14 kbp) were identified in 20 isolates of *Campylobacter lari* (urease-negative *C. lari* [*n*=7]; urease-positive thermophilic *Campylobacter* [*n*=13]). The nucleotide sequences of an approximately 70-bp non-coding region, including the typical promoter structure, showed sequence differences at 12 loci among 21 isolates including *C. lari* RM2100. The putative σ^{70} promoter region upstream of the putative open reading frame (ORF), a start codon TTG and a probable ribosome binding site, AGGA, for the *vacJ* gene were also identified in all 21 *C. lari* isolates examined. Each ORF for the *vacJ* terminated with a TAA stop codon. No hypothetical transcriptional terminators were identified within the amplicons. The putative ORFs of the *vacJ* gene from 21 *C. lari* isolates consisted of 684 bases, similarly differing from those of the other thermophilic campylobacters (696 bases for *C. jejuni* RM1221 and NCTC11168 and *C. coli* RM2228; 690 for *C. upsaliensis* RM3195). Reverse transcription PCR analysis confirmed the transcription of the *vacJ* gene in the *C. lari* cells. A neighbour joining tree suggested a strong molecular discrimination efficacy between UPTC and UN *C. lari* employing *vacJ* nucleotide sequence information. The *vacJ* gene homologue from *C. lari* organisms appears not to be a lipoprotein signal peptide or a signal peptide *in silico*.

KEY WORDS: Amino acid sequence.
Campylobacter lari.
Nucleotides.
vacJ gene.

Introduction

Campylobacter species, primarily *C. jejuni* and *C. coli*, are curved Gram-negative bacteria that are a recognised cause of bacterial diarrhoea worldwide. Moreover, they lead to enteric disease in several species of domestic animals.

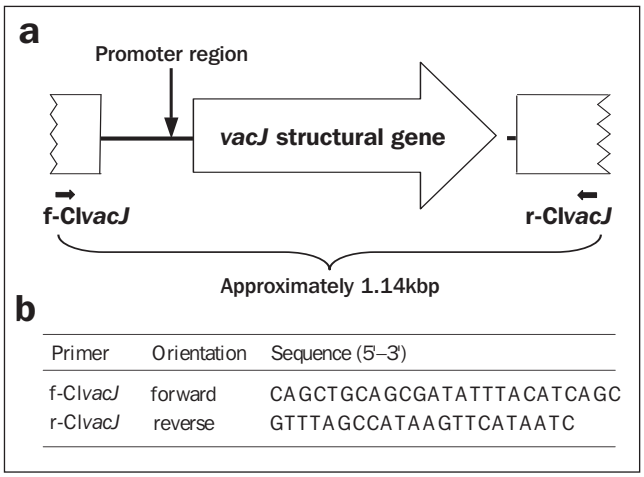
In 2000, the genome sequence of *C. jejuni* NCTC11168 was described by Parkhill *et al*.[1] More recently, in 2005, a comparison of five *Campylobacter* genomes was reported by Fouts *et al*.[2] and it was suggested that different *Campylobacter* species would possess different pathogenic mechanisms in human infection.

In addition, *C. lari* is a thermophilic *Campylobacter* species that is resistant to nalidixic acid, one of the compounds generally used to discriminate this species from *C. jejuni* and *C. coli*.[3,4] *C. lari* organisms were first isolated from mammalian and avian species, particularly seagulls of the genus *Larus*,[3,4] and has been shown to be a cause of clinical infection.[5,6] Although, to date, cases of human illness associated with *C. lari* have been infrequent, clinical isolates from patients have added to more than 110 isolates from 30 cases in several countries over the last 25 years. Thus, *C. lari* may be a severe and potential pathogenic agent for humans.

An atypical group of isolates of urease-positive thermophilic *Campylobacter* (UPTC) was first isolated from the natural environment in England in 1985.[7] Thereafter, this organism was described as a biovar or variant of *C. lari*.[8,9] Although subsequent reports verified four human isolates from the faeces of two adults with diarrhoeal disease, from the appendix of a child with appendicitis and from the urine of a patient with a urinary tract infection in France,[8,10] the possible association of UPTC with human disease remains unclear. Additional isolates of UPTC have also been reported in Northern Ireland,[11–13] in The Netherlands[14] and in Japan.[15,16] Until now, only these four clinical cases of UPTC isolates in France have been published. Thus, these two representative taxa, namely urease-negative (UN) *C. lari* and UPTC occur within the species of *C. lari*.[17]

A putative surface-exposed lipoprotein encoded on the virulence-associated chromosomal J (*vacJ*) gene, required for

Correspondence to: Dr. 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



b

Primer	Orientation	Sequence (5′–3′)
f-Cl*vacJ*	forward	CAGCTGCAGCGATATTTACATCAGC
r-Cl*vacJ*	reverse	GTTTAGCCATAAGTTCATAATC

Fig. 1. A schematic representation of full-length *vacJ* structural gene homologue of *C. lari*, including locations of the primer pair (f-Cl*vacJ*/r-Cl*vacJ*), based on sequence information of the *vacJ* gene and its adjacent genetic loci in *C. lari* RM2100 (A) and nucleotide sequences of the newly constructed primer pair for PCR amplification of the full-length *vacJ* gene (B).

intercellular spreading in *Shigella flexneri,* has been demonstrated.[18] In *S. flexneri*, cloning and sequencing of the *vacJ* region indicated that the *vacJ* gene encoded a 28 kDa protein, possessing a signal peptide at the N-terminus, which contained the motif characteristic of lipoproteins.[18] In addition, recently, Santos *et al*. described one (VacJ) of the 68 proteins increased in the *Pseudomonas putida* KT2440 response to phenol-induced stress by quantitative proteomics.[19]

In relation to the *vacJ* of thermophilic *Campylobacter*, *vacJ* homologues encoding VacJ lipoprotein or VacJ-like protein recently were identified in *C. jejuni* (subsp. *jejuni*) NCTC11168 (DDBJ/EMBL/GenBank accession number: NC_002163), *C. jejuni* RM1221 (NC_003912), *C. coli* RM2228 (AAFL01000008), *C. upsaliensis* RM3195 (AAFJ01000005) and *C. lari* RM2100 (AAFK01000002) following whole-genome shotgun sequencing analysis, but such reports have yet to appear.

Thus, the *vacJ* gene homologue is distributed among thermophilic *Campylobacter* organisms. Therefore, it would be worthwhile to clarify whether or not the *vacJ* homologue is a genomic gene essential for pathogenicity to spread intercellularly in thermophilic *Campylobacter* organisms.

This study aims to clone, sequence, analyse and characterise the *vacJ* gene homologue sequences, including the promoter region, in isolates of *C. lari* comprising UN *C. lari* and UPTC, and then compare their characteristics to each other with other thermophilic campylobacters in order to clarify the presence of this *vacJ* gene homologue in *C. lari* organisms. In addition, the study aims to confirm the expression of the gene in *C. lari* cells.

Materials and methods

The bacterial isolates, UN *C. lari* (n=7) and UPTC (n=13), used in the study are shown in Table 1. These *C. lari* isolates were cultured on Butzler medium at 37°C for 48 h in an aerobic jar on blood agar base No. 2 (Oxoid) 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). Genomic DNA was prepared using sulphate and proteinase K, phenol-chloroform extraction and ethanol precipitation.[20]

```
                                T-rich region          -10 region              SD     Start codon
C.lari JCM2530T      201  TGTAAAAATTTATTTTAATTTTTGTATAATCTATGATTTTAATTTATGAAATTTTAAAAGGAAAAA-CTTGTTAA  274
C.lari 175           201  ............................T.....................................-........  274
C.lari 296           201  ............................T.....................................-........  274
C.lari 298           201  ............................T.....................................-........  274
C.lari 48            201  ............................T..A..................................-........  274
C.lari 84C-1         201  ............................T.....................................-........  274
C.lari 448           201  ............................T.....................................-........  274
UPTC NCTC12892       201  ............................T..C..................................-........  274
UPTC NCTC12893       201  ............................T..C..................................-........  274
UPTC NCTC12894       201  ............................T..C.....C............................-........  274
UPTC NCTC12895       201  ............................T..C.....C............................-........  274
UPTC NCTC12896       200  ............................T..C.....C............................-........  273
UPTC 89049           192  ............................A..A.....C...........................G........  266
UPTC 92251           201  ............................T..A.G...C.....G.....................-........  274
UPTC CF89-12         201  ...........................GAA...C.......A.G....-...-.......G........  273
UPTC CF89-14         201  ...........................GAA...C.......A.G....-...-.......G........  273
UPTC 99              202  ............................T..C.....................................-........  275
UPTC A1              192  ............................A..A.....C...........................G........  266
UPTC A2              192  ............................A..A.....C...........................G........  266
UPTC A3              192  ............................A..A.....C...........................G........  266
C.lari RM2100        201  ............................T.....................................-........  274
                          ****************************** *    *** ***** * * **** *** ******* ********
```

Takaku *et al*.

Fig. 2. Nucleotide sequence alignment analyses of an approximately 70-bp NC region, including the typical promoter structure of the 21 *C. lari* isolates, including *C. lari* RM2100. Dots indicate identical bases; changes are so indicated; dashes are deletions; positions identical in all isolates are marked by asterisks; numbers at the left and right refer to base pairs of the *vacJ* gene sequence of the isolates, respectively. The sequence data of the *vacJ* sequence homologues determined in the present study are accessible in the DDBJ/EMBL/GenBank, as described in Table 1.

Table 1. Isolates of thermophilic *Campylobacter* used in the present study, some reference strains, accession numbers of the nucleotide sequence data of approximately 1.14 kbp region containing full-length *vacJ* structural gene accessible in DDBJ/EMBL/GenBank and putative ORF and CMW of the gene

Isolate	Source	Country	ORF	Numbers of amino acids	CMW (Da)	Accession number
UN *C. lari* JCM2530[T]	Seagull	Japan	684	228	26,400	AB433597
UN *C. lari* 175	Black-tailed gull	Japan	684	228	26,399	AB433598
UN *C. lari* 296	Human	Canada	684	228	26,400	AB433599
UN *C. lari* 298	Human	Japan	684	228	26,400	AB433600
UN *C. lari* 48	Mussel	Northern Ireland	684	228	26,414	AB433601
UN *C. lari* 84C-1	Human	Northern Ireland	684	228	26,399	AB433602
UN *C. lari* 448	Mussel	Northern Ireland	684	228	26,399	AB433603
UPTC NCTC12892	River water	England	684	228	26,416	AB433604
UPTC NCTC12893	River water	England	684	228	26,416	AB433605
UPTC NCTC12894	Sea water	England	684	228	26,484	AB433606
UPTC NCTC12895	Mussel	England	684	228	26,510	AB433607
UPTC NCTC12896	Mussel	England	684	228	26,420	AB433608
UPTC 89049	Human	France	684	228	26,546	AB433609
UPTC 92251	Human	France	684	228	26,464	AB433610
UPTC CF89-12	River water	Japan	684	228	26,528	AB433611
UPTC CF89-14	River water	Japan	684	228	26,528	AB433612
UPTC 99	Sea water	Northern Ireland	684	228	26,464	AB433613
UPTC A1	Seagull	Northern Ireland	684	228	26,554	AB433614
UPTC A2	Seagull	Northern Ireland	684	228	26,554	AB433615
UPTC A3	Seagull	Northern Ireland	684	228	26,554	AB433616
C. lari RM2100	Human	USA	684	228	26,433	AAFK01000002
C. jejuni RM1221	Chicken	USA	696	232	26,475	NC_003912
C. jejuni NCTC 11168	Human	UK	696	232	26,507	NC_002163
C. coli RM2228	Chicken	USA	696	232	26,386	AAFL01000008
C. upsaliensis RM3195	Human	USA	690	230	26,091	AAFJ01000005

ORF: open reading frame

CMW: calculated molecular weights

JCM: Japan Collection of Microorganisms.

A polymerase chain reaction (PCR) primer pair (f-Cl*vacJ* and r-Cl*vacJ*) was devised *in silico* to amplify the full-length *vacJ* gene, including the promoter region (approximately 750 base pair [bp]) within *C. lari*, based on the sequence information of the *vacJ* structural gene and its adjacent genetic loci in *C. lari* RM2100 (AAFK01000002),[2] as shown in Figure 1.

The PCR mixture contained 100 ng template DNA, 10 mmol/L Tris-HCl (pH 8.3), 50 mmol/L KCl, 1.5 mmol/L $MgCl_2$, 100 µmol/L each dNTP, 30 pmol/L each primer, and 1 unit EX *Thermus aquaticus* (*Taq*) DNA polymerase (Takara Bio, Shiga, Japan). The PCR was performed in a 50 µL volume at 94˚C for 3 min, 30 cycles at 94˚C for 1 min, 50˚C for 0.5 min, 72˚C for 1.5 min, and finally 72˚C for 7 min.

Amplified PCR products were separated by 1% (w/v) agarose gel electrophoresis in 0.5xTBE, then extracted and purified from the gel, as described previously.[21,22] Purified PCR products were then cloned into pGEM-T vector (Promega, Tokyo, Japan) using the TA cloning procedure. Following nucleotide sequencing with M13 (Hitachi SQ5500-EL DNA autosequencer), sequence analysis was carried out using the Genetyx-Mac (version 9; Genetyx, Tokyo, Japan) computer software. For accurate sequencing, multiple TA-cloned PCR products were sequenced, and those corresponding to the PCR primers of f-Cl*vacJ* and r-Cl*vacJ* were excluded from the sequence data of the approximately 1.14 kbp region containing the full-length *vacJ* structural gene accessible in DDBJ/EMBL/GenBank.

Total cellular RNA was extracted and purified from *C. lari* cells using RNAprotect Bacteria Reagent and the RNeasy Mini Kit (Qiagen, Tokyo, Japan). Reverse transcription (RT)-PCR analysis was carried out with a primer pair (f-/r-Cl*vacJ*mid; f-Cl*vacJ*mid, 5′-GGHGARGAATTTRRAAGATTT-3′ and r-Cl*vacJ*mid, 5′-TKGCGTCTTTGYTCRTAHGCATC-3′) using the Takara One-Step RNA PCR kit (AMV; Takara Bio). This primer pair was expected to generate a product of the *vacJ* gene segment of approximately 370 bp.

Nucleotide and putative amino acid sequences of approximately 680 bp of the full-length *vacJ* structural gene from the 20 *C. lari* isolates were compared and with the accessible sequence data of other thermophilic campylobacters using CLUSTAL W software,[23] incorporated in DDBJ. Finally, a phylogenetic tree was constructed by the neighbour joining (NJ) method.[24]

Table 2. Sequence similarities (%) of the nucleotides (upper right) and amino acids (lower left) of the full-length *vacJ* structural gene from the 21 *C. lari* isolates and other thermophilic campylobacters.

		1	2	3	4	5	6	7	8	9	10	11
1	*C. lari* JCM2530T		99.3	100.0	99.9	99.7	99.27	99.3	91.4	91.7	91.7	91.8
2	*C. lari* 175	99.6		99.3	99.4	99.3	100.0	100.0	91.1	91.4	91.4	91.5
3	*C. lari* 296	100.0	99.6		99.9	99.7	99.3	99.3	91.4	91.7	91.7	91.8
4	*C. lari* 298	100.0	99.6	100.0		99.9	99.4	99.4	91.5	91.8	91.8	92.0
5	*C. lari* 48	99.6	99.1	99.6	99.6		99.3	99.3	91.4	91.7	91.7	92.1
6	*C. lari* 84C-1	99.6	100.0	99.6	99.6	99.1		100.0	91.1	91.4	91.4	91.5
7	*C. lari* 448	99.6	100.0	99.6	99.6	99.1	100.0		91.1	91.4	91.4	91.5
8	UPTC NCTC12892	92.5	92.1	92.5	92.5	92.1	92.1	92.1		99.7	95.8	95.9
9	UPTC NCTC12893	92.5	92.1	92.5	92.5	92.1	92.1	92.1	100.0		95.8	95.9
10	UPTC NCTC12894	92.5	92.1	92.5	92.5	92.1	92.1	92.1	97.4	97.4		98.4
11	UPTC NCTC12895	92.5	92.1	92.5	92.5	93.0	92.1	92.1	97.4	97.4	98.3	
12	UPTC NCTC12896	92.5	92.1	92.5	92.5	92.1	92.1	92.1	97.4	97.4	99.1	98.3
13	UPTC 89049	92.5	92.1	92.5	92.5	93.0	92.1	92.1	96.5	96.5	97.4	98.3
14	UPTC 92251	92.5	92.1	92.5	92.5	92.1	92.1	92.1	97.4	97.4	99.1	98.3
15	UPTC CF89-12	93.0	92.5	93.0	93.0	92.5	92.5	92.5	96.1	96.1	96.9	96.9
16	UPTC CF89-14	93.0	92.5	93.0	93.0	92.5	92.5	92.5	96.1	96.1	96.9	96.9
17	UPTC 99	92.5	92.1	92.5	92.5	93.0	92.1	92.1	97.4	97.4	99.1	99.1
18	UPTC A1	91.7	91.2	91.7	91.7	92.1	91.2	91.2	95.6	95.6	97.4	97.4
19	UPTC A2	91.7	91.2	91.7	91.7	92.1	91.2	91.2	95.6	95.6	97.4	97.4
20	UPTC A3	91.7	91.2	91.7	91.7	92.1	91.2	91.2	95.6	95.6	97.4	97.4
21	*C. lari* RM2100	99.1	99.6	99.1	99.1	98.7	99.6	99.6	91.7	91.7	91.7	91.7
22	*C. jejuni* RM1221	50.0	50.4	50.0	50.0	50.4	50.4	50.4	50.2	50.2	50.4	50.9
23	*C. jejuni* NCTC11168	49.8	50.0	49.8	49.8	50.2	50.0	50.0	49.6	49.6	50.2	50.6
24	*C. coli* RM2228	52.8	52.4	52.8	52.8	53.2	52.4	52.4	52.2	52.2	51.7	51.7
25	*C. upsaliensis* RM3195	48.5	48.5	48.5	48.5	48.9	48.5	48.5	48.1	48.1	48.5	48.5

Results

Following construction of a PCR primer pair (f-/r-Cl*vacJ*) for amplification of the *vacJ* gene sequence, including the promoter region from the *C. lari* isolates, approximately 750 bp of promoter and structural gene regions for *vacJ* and its adjacent genetic loci (approximately 1.14 kbp) were successfully amplified, TA-cloned and sequenced from all 20 isolates of *C. lari* employed.

When the nucleotide sequence multiple-alignment analyses were carried out, nucleotide sequences of approximately 1.14 kbp obtained from all the isolates identified a partial and a putative molybdenum ABC transporter, periplasmic molybdate-binding protein (*modA*) gene and its promoter region, a putative promoter region for the *vacJ* gene and *vacJ* gene structural gene homologue, and a partial and putative periplasmic protein Cj1372 (*CLA*0409) gene.

The nucleotide sequences of the approximately 70-bp non-coding (NC) region, including the typical promoter structure (nucleotide positions 201–270 bp for UN *C. lari* JCM2530^T) of *C. lari* analysed in the present study showed nucleotide sequence differences at the 12 loci among the 21 *C. lari* isolates including the *C. lari* RM2100 strain (Fig. 2). The nucleotide positions used were for those of UN *C. lari* JCM2530^T (AB433597).

Based on the sequence information of the *vacJ* gene homologue from *C. lari* RM2100 (AAFK01000002), a –10 region (TTTAAT, 238-243 bp) was identified as the typical transcriptional promoter structure for the *C. lari* isolates upstream of the putative open reading frame (ORF) of the *vacJ* structural gene homologue (268–951 bp), as well as the start codon TTG (268-270 bp) (Fig. 2). However, no consensus sequences at the –35 region were identified, and a semi-conserved T-rich region (209–223 bp; T=12/15) was identified instead of the region, as shown in *Rpo*D promoters in the genome of *C. jejuni*.[25] Therefore, the *vacJ* gene homologue may be transcribed by the σ^{70} factor in *C. lari* organisms. Thus, the putative σ^{70} transcriptional promoter structures were identified for the *vacJ* structural gene homologue in all 21 isolates, including the *C. lari* RM2100 strain, examined (Fig. 2). The TTG start codon was also identified for the *vacJ* gene homologue in all 20 isolates examined. A probable ribosome binding (RB) site (Shine-Dalgarno [SD] sequence)[26] complementary to a highly conserved sequence of CCUCCU close to the 3′ end of 16S ribosomal RNA, AGGA (259–262 bp), was also identified for *vacJ*.

Each putative ORF for the *vacJ* gene homologue of all 21 *C. lari* isolates terminated with a TAA stop codon (952–954 bp). In addition, as shown in Table 1, the putative ORFs of the *vacJ* gene homologue from the 21 *C. lari* isolates consisted of

12	13	14	15	16	17	18	19	20	21	22	23	24	25
91.5	91.4	91.5	91.1	91.1	91.5	91.1	91.1	91.5	99.1	59.9	60.2	58.6	57.5
91.2	91.1	91.2	90.8	90.8	91.2	90.8	90.8	91.2	99.9	60.5	60.5	58.5	57.3
91.5	91.4	91.5	91.1	91.1	91.5	91.1	91.1	91.5	99.1	59.9	60.2	58.6	57.5
91.7	91.5	91.7	91.2	91.1	91.7	91.2	91.2	91.7	99.3	59.8	60.1	58.8	57.7
91.5	91.7	91.5	91.1	91.1	91.8	91.4	91.4	91.8	99.1	59.9	60.2	58.9	57.8
91.2	91.1	91.2	90.8	90.8	91.2	90.8	90.8	91.2	99.9	60.5	60.5	58.5	57.3
91.2	91.1	91.2	90.8	90.8	91.2	90.8	90.8	91.2	99.9	60.5	60.5	58.5	57.3
95.9	95.6	95.9	94.6	94.6	96.5	94.6	94.6	94.4	90.9	60.9	60.5	58.8	58.7
95.9	95.6	95.9	94.3	94.3	96.5	94.6	94.6	94.4	91.2	61.1	60.6	58.9	58.7
99.0	96.1	97.7	95.6	95.6	98.7	95.3	95.3	95.5	91.2	59.9	59.9	58.3	57.4
98.5	96.5	97.5	95.5	95.5	98.5	95.5	95.5	95.6	91.4	60.1	60.2	58.1	58.1
	96.1	98.4	95.6	95.6	99.1	95.0	95.0	95.2	91.1	59.8	59.9	58.2	57.4
97.4		95.2	96.9	96.9	96.4	98.7	98.7	98.5	90.9	60.1	60.2	58.8	57.7
99.1	97.4		94.7	94.7	98.7	94.2	94.2	94.3	91.1	59.2	59.3	58.5	57.1
96.9	97.8	96.9		100.0	95.6	96.8	96.8	96.4	90.6	60.2	60.5	58.9	58.4
96.9	97.8	96.9	100.0		95.6	96.8	96.8	96.4	90.6	60.2	60.5	58.9	58.4
99.1	98.3	99.1	96.9	96.9		95.3	95.3	95.5	91.1	60.1	60.1	58.5	57.7
96.5	99.1	96.5	97.8	97.8	97.4		100.0	99.3	90.6	60.3	60.2	58.9	58.0
96.5	99.1	96.5	97.8	97.8	97.4	100.0		99.3	90.6	60.3	60.2	58.9	58.0
96.5	99.1	96.5	97.8	97.8	97.4	100.0	100.0		91.1	60.2	60.1	58.9	57.8
91.7	91.7	91.7	92.1	92.1	91.7	90.8	90.8	90.8		60.3	60.3	58.3	57.1
50.4	50.9	50.4	50.0	50.0	50.9	50.9	50.9	50.9	50.0		97.4	81.2	70.0
50.2	50.2	50.2	49.4	49.4	50.6	50.2	50.2	50.2	49.6	97.0		81.9	70.1
51.3	52.2	51.7	51.3	51.3	52.2	52.2	52.2	52.2	51.9	87.5	88.8		70.7
48.5	48.9	48.1	48.1	48.1	48.9	48.9	48.9	48.9	48.1	70.3	70.3	70.3	

684 bases (228 amino acid residues), similarly differing from those of the other thermophilic campylobacters (696 [232] for *C. jejuni* RM1221 and NCTC11168 and *C. coli* RM2228; 690 [230] for *C. upsaliensis* RM3195).

In the region upstream of the promoter for the *vacJ* structural gene homologue, an NC region of approximately 150 bp (49–200 bp) and a partial and putative ORF for the *modA* gene (1–48 bp) occurred in the reverse direction to the *vacJ* gene homologue. In the NC region (approximately 150 bp), a putative promoter structure at the –10 (TATAAT; 86–81 bp) was also identified. In the partial and putative ORF (48 bp), eight polymorphic sites occurred among the 21 *C. lari* isolates. For the partial and putative ORF, an RB site (AGGA; 58–55 bp) and an ATG start codon (46–48 bp) were identified in all 21 isolates.

In the region downstream of the *vacJ* structural gene homologue, an NC region of approximately 9 bp (955–963 bp) and a partial and putative ORF for *CLA*0409 (964–1140 bp) occurred.

Regarding the transcriptional terminator structure, hypothetical terminators were not identified between the NC region and the partial and putative ORF within the 21 *C. lari* isolates examined. Therefore, a *vacJ* structural gene homologue might construct a polycistronic operon together with the *CLA*0409 and any other structural gene(s) downstream.

When RT-PCR analysis was carried out for the RNA components extracted and purified from the *C. lari* cells with the primer pair (f-/r-Cl*vacJ*mid), a positive RT-PCR signal (approximately 370 bp) was detected at around the 400-bp region with UN *C. lari* JCM2530^T (Fig. 3, Lane 1) and UPTC CF89-12 (Lane 4). When the positive RT-PCR products from the two isolates were sequenced, both were identified to be *vacJ* (data not shown). Thus, the transcription of the *vacJ* gene was confirmed in the *C. lari* isolate cells. As an approximate 370 bp fragment was also observed with genomic DNA as a PCR template in both isolates (Fig. 3, Lanes 3 and 6), this primer pair of f-/r-*vacJ*mid may be useful for the detection of approximately 370-bp fragments of *vacJ* gene from *C. lari* isolates.

An NJ tree was constructed based on the present nucleotide sequence information of the *vacJ* gene homologue of the *C. lari* isolates and the other thermophilic *Campylobacter* reference strains accessible in the DDBJ/EMBL/GenBank (Fig. 4). There are some major clusters in the NJ dendrogram constructed and the phylogenetic tree demonstrated that the 20 *C. lari* isolates and a reference UN *C. lari* RM2100 strain formed some minor clusters with genetic variability, separating them from the other thermophilic campylobacters (*C. jejuni*, *C. coli* and *C. upsaliensis*; Fig. 4). In addition, the phylogenetic tree also demonstrated that the 21 *C. lari* isolates formed two

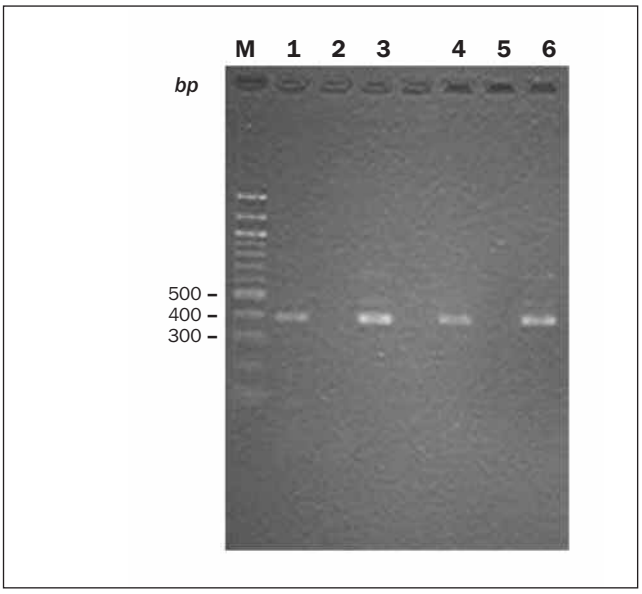


Fig. 3. RT-PCR analysis of the *vacJ* structural gene transcript expressed in the *C. lari* isolates. Lane M: 100-bp DNA ladder, lanes 1–3: UN *C. lari* JCM2530^T, lanes 4–6: UPTC CF89-12. Lanes 1 and 4: with total cellular RNA and AMV reverse transcriptase (RTase) XL, lanes 2 and 5: with total cellular RNA but without AMVRTase XL, lanes 3 and 6: without total cellular RNA but with the genomic DNA.

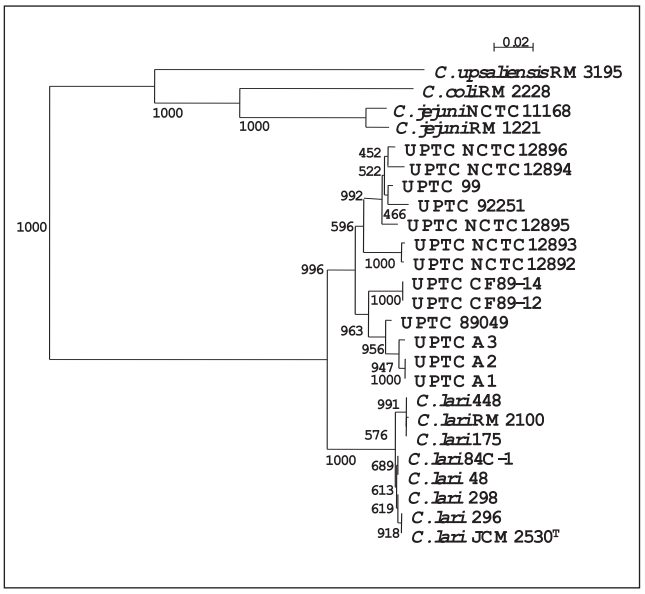


Fig. 4. A phylogenetic tree constructed by the NJ method, based on the nucleotide sequence similarity data of the full-length *vacJ* structural gene homologue from *C. lari* isolates and three other thermophilic campylobacters. Out group is *C. upsaliensis* RM3195. Boot-strap values of 1000 are shown at the branch point. Bar, 0.02 (evolutionary distance).

distinctly different clusters consisting of eight UN *C. lari* and 13 UPTC isolates.

Discussion

This is the first description of the construction of a PCR primer pair for amplification of the *vacJ* gene homologue, including the promoter region, its successful PCR amplification and TA-cloning and sequencing with 20 *C. lari* isolates.

In the present study, nucleotide sequences of the full-length *vacJ* structural gene homologue from 21 *C. lari* isolates including *C. lari* RM2100 showed 90.6–100% sequence similarity, and approximately 57.1–61.1% similarity to those of the four *C. jejuni*, *C. coli* and *C. upsaliensis* strains. Approximately 90.6–92% nucleotide sequence similarity was also shown between the eight UN *C. lari* and 13 UPTC isolates (Table 2). In addition, the putative ORFs of the *vacJ* gene homologue from 21 *C. lari* isolates showed 90.8–100% amino acid sequence similarity, and approximately 48.1–53.2% similarity to those of the four strains.

Approximately 90.8–93% amino acid sequence similarity was shown between the UN *C. lari* and UPTC isolates (Table 2). These suggest that the 21 *C. lari* isolates share higher *vacJ* gene homologue sequence similarities of nucleotide and amino acid levels, compared to the other thermophilic campylobacters.

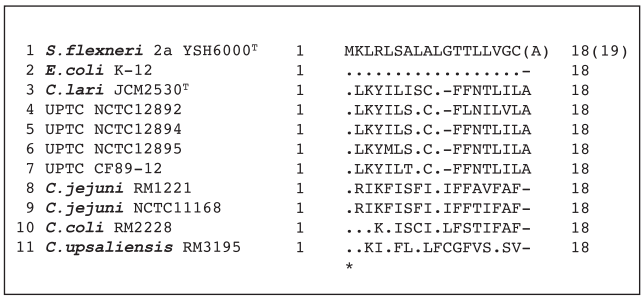
The authors have already reported that the partial and putative ORFs of the cytolethal distending toxin B (*cdtB*) gene fragments (equivalent to 90% segment of the *cdtB* structural gene of *C. lari*) from 25 *C. lari* isolates including *C. lari* RM2100 showed 78.2–99.6% amino acid sequence similarity, and 63.9–88% nucleotide and 66.6–78.5% amino acid sequence similarity to those of the ORFs of *C. jejuni* RM1221 and *C. coli* RM2228 reference strains.[27] Thus, interestingly, lower genetic heterogeneity of nucleotide and amino acid sequences to those of the ORFs of *C. jejuni* and *C. coli* strains of the *vacJ* gene homologue than those of the partial *cdtB* gene fragments was demonstrated among the *C. lari* isolates consisting of UN *C. lari* and UPTC. In addition, higher genetic heterogeneity of the *vacJ* gene homologue than those of the partial *cdtB* gene fragment occurred among the three thermophilic campylobacters *C. lari*, *C. jejuni* and *C. coli*.

Thus, in the present study, the *vacJ* gene homologue appears to be relatively homologous in both nucleotide and deduced amino acid sequence levels among the *C. lari* isolates, but it has an extremely high sequence variability against two other thermophilic campylobacters (*C. jejuni* and *C. coli*).

As shown in Figure 4, the NJ tree indicated that the nucleotide sequence information of the *vacJ* gene homologue showed strong molecular discrimination efficacy between UPTC and UN *C. lari* organisms. The present phylogenetic analysis using the NJ method may suggest that nucleotide sequence information of the *vacJ* gene homologue could be a useful epidemiological tool in the thermophilic *Campylobacter* organisms.

The putative ORFs of the *vacJ* gene homologue from the 21 *C. lari* isolates were identified to consist of similar 684 bp (228 amino acid) residues, but differed from those of the other three thermophilic *Campylobacter* species. In addition, this deduced amino acid sequence, consisting of 228 amino acid residues for the *vacJ* gene homologue of *C. lari*, was also much shorter than those (251) of the *S. flexneri* 2a YSH6000^T strain,[18] *Escherichia coli* K-12 (NC_000913) and *Salmonella enterica* Typhi CT18 (NC_003198).

Figure 5 shows sequence alignments of the first 18 amino acid residues containing the motif regions characteristic of the lipoproteins at the N-terminus[18] in the putative ORF of



1	***S.flexneri*** 2a YSH6000^T	1	MKLRLSALALGTTLLVGC(A)	18(19)
2	***E.coli*** K-12	1	..................-	18
3	***C.lari*** JCM2530^T	1	.LKYILISC.-FFNTLILA	18
4	UPTC NCTC12892	1	.LKYILS.C.-FLNILVLA	18
5	UPTC NCTC12894	1	.LKYILS.C.-FFNTLILA	18
6	UPTC NCTC12895	1	.LKYMLS.C.-FFNTLILA	18
7	UPTC CF89-12	1	.LKYILT.C.-FFNTLILA	18
8	***C.jejuni*** RM1221	1	.RIKFISFI.IFFAVFAF-	18
9	***C.jejuni*** NCTC11168	1	.RIKFISFI.IFFTIFAF-	18
10	***C.coli*** RM2228	1	...K.ISCI.LFSTIFAF-	18
11	***C.upsaliensis*** RM3195	1	..KI.FL.LFCGFVS.SV- *	18

Fig. 5. Partial amino acid sequence alignment analyses of the motif regions characteristic of lipoproteins at the N-terminus in the putative ORF of the *vacJ* structural gene homologue from the 21 *C. lari* isolates examined. The corresponding regions of the putative ORF from some thermophilic *Campylobacter* reference strains, *S. flexneri*[19] and *E. coli* K-12 (NC_000913) were also aligned for comparison. Numbers on the left and right refer to the amino acid positions in each putative ORF for the *vacJ* from the isolates. For others, refer to the legend to Figure 2. Sequences of the first 18 amino acid residues in the UN *C. lari* JCM2530^T (3) were identical to those of *C. lari* 175, 296, 298, 48, 84C-1, 448 and RM2100; UPTC NCTC12892 (4), NCTC12893; UPTC NCTC12894 (5), NCTC12896, 92251 and UPTC99; UPTC CF89-12 (7), 89049, CF89-14, A1, A2 and A3.

the *vacJ* structural gene homologue from the *C. lari* isolates, as well as those of *Shigella flexneri* (D06293)[18] and *E. coli* K-12 (NC_000913). The 18 amino acid residues identified in *S. flexneri* are perfectly aligned and fully conserved in the *E. coli* K-12 used for comparison. Surprisingly, however, the first common 18 amino acid residues in *S. flexneri*[18] and *E. coli* K-12 (NC_000913) are not conserved among the residues of the 25 thermophilic *Campylobacter* organisms (Fig. 5).

Regarding the VacJ of *S. flexneri*, the first 18 amino acid residues were those of a typical signal sequence with a predicted peptidase cleavage site between Gly-17 and Cys-18.[18] In addition, the motif Leu-X-Gly-Cys is characteristic of the processing site for lipoproteins.[18,28] No typical signal sequence and motif Leu-X-Gly-Cys sequence characteristics of the processing site for lipoproteins were detected in the *vacJ* gene homologues from all 25 thermophilic *Campylobacter* organisms examined (Fig. 5).

Finally, the study attempted to discover whether or not the *Campylobacter vacJ* gene homologues examined were lipoprotein using the lipoprotein prediction program (CBS Prediction Servers) *in silico*. When the first 18 amino acid residues of the *vacJ* from *S. flexneri* and *E. coli* were examined, lipoprotein signal peptide (SpII) score was shown to be 20.7085 (margin=9.3331; cleavage=17–18; Post 2=A). The *vacJ* resulted in encoding outer membrane lipoprotein. However, with *C. jejuni*, *C. coli* and *C. upsaliensis*, signal peptide (SpI) scores were shown to be 10–14 (margin=11–14; cleavage=19–20). Consequently, secretory protein was a candidate for the three thermophilic *Campylobacter* species (*C. jejuni*, *C. coli* and *C. upsaliensis*). In addition, as both scores of SpI and SpII were very low for *C. lari*, the *vacJ* gene homologue from the *C. lari* orgnisms appears not to be lipoprotein signal nor signal peptides.

In conclusion, this study showed that the *vacJ* gene homologue was transcribed in *C. lari* cells. Therefore, elucidation of the functions of the *vacJ* gene homologue will lead to a full understanding of its roles in *C. lari*. □

This research was partially supported 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) (20580346) from the Ministry of Education, Culture, Sports, Science and Technology of Japan (to MM). MM and JEM were funded through a Great Britain Sasakawa Foundation (Butterfield) Award to examine the clinical significance of Campylobacter *infection in the UK and Japan.*

References

1 Parkhill J, Wren BW, Mungall K *et al*. The genome sequence of the food-borne pathogen *Campylobacter jejuni* reveals hypervariable sequences. *Nature* 2000; **403**: 665–8.
2 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.
3 Skirrow MB, Benjamin J. '1001' campylobacters: cultural characteristics of intestinal campylobacters from man and animals. *J Hyg (Camb)* 1980; **85**: 427–42.
4 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.
5 Nachamkin I, Stowell C, Skalina D *et al*. *Campylobacter laridis* causing bacteremia in an immunosuppressed patient. *Ann Int Med* 1984; **101**: 55–7.
6 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.
7 Bolton FJ, Holt AV, Hutchinson DN. Urease-positive thermophilic campylobacters. *Lancet* 1985; **i**: 1217–8.
8 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* 1988; **26**: 1050–1.
9 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.
10 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.
11 Wilson IG, Moore JE. Presence of *Salmonella* spp. and *Campylobacter* spp. in shellfish. *Epidemiol Infect* 1996; **116**: 147–53.
12 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.
13 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.
14 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.
15 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.

16 Matsuda M, Shibuya T, Itoh Y *et al*. First isolation of urease-positive thermophilic *Campylobacter* (UPTC) from crows (*Coruvs levaillantii*) in Japan. *Int J Hyg Environ Health* 2002; **205**: 321–4.
17 Matsuda M, Moore JE. Urease-positive thermophilic *Campylobacter* species. *Appl Environ Microbiol* 2004; **70**: 4415–8.
18 Suzuki T, Murai T, Fukuda I *et al*. Identification and characterization of a chromosomal virulence gene, *vacJ*, required for intercellular spreading of *Shigella flexneri*. *Mol Microbiol* 1994; **11**: 31–41.
19 Santos PM, Benndorf D, Sa-Correia I. Insights into *Pseudomonas putida* KT2440 response to phenol-induced stress by quantitative proteomics. *Proteomics* 2004; **4**: 2640–52.
20 Sambrook J, Russell DW. *Molecular cloning; a laboratory manual* 3rd edn. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, 2001.
21 Sekizuka T, Seki K, Hayakawa T *et al*. Phenotypic characterisation of flagellin and flagella of urease-positive thermophilic campylobacters. *Br J Biomed Sci* 2004; 61: 186–9.
22 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 Leeuwen* 2005; **88**: 113–20.
23 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; **212**: 4673–80.
24 Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 1987: **4**: 406–25.
25 Petersen L, Larsen TS, Ussery DX *et al*. *Rpo*D promoters in *Campylobacter jejuni* exhibit a strong periodic signal instead of a –35 box. *J Mol Biol* 2003; **326**; 1361–72.
26 Benjamin L. *Genes VII.* Oxford: Oxford University Press, 2000.
27 Shigematsu M, Harada Y, Sekizuka T *et al*. Genetic heterogeneity of the cytolethal distending toxin B (*cdt*B) gene locus among isolates of *Campylobacer lari*. *Br J Biomed Sci* 2006; **63**: 179–81.
28 Wu H C, Tokunaga M. Biogenesis of lipoproteins in bacteria. *Curr Top Microbiol Immunol* 1986; **125**: 127–57.