# Cloning and structural analysis of the full-length cytolethal distending toxin (*cdt*) gene operon from *Campylobacter lari*

M. MATSUDA*, M. SHIGEMATSU*, A. TAZUMI*, T. SEKIZUKA*, S. TAKAMIYA†, B. C. MILLAR‡, I. TANEIKE* and J. E. MOORE‡

**Laboratory of Molecular Biology, Graduate School of Environmental Health Sciences, Azabu University, Fuchinobe 1-17-71, Sagamihara; †Department of Molecular and Cellular Parasitology, Juntendo University School of Medicine, Hongo 2-1-1, Bunkyoku, Tokyo 113-8421, Japan; and ‡Department of Bacteriology, Northern Ireland Public Health Laboratory, Belfast City Hospital, Belfast BT9 7AD, Northern Ireland, UK*

*Accepted: 14 March 2008*

**ABSTRACT**

Polymerase chain reaction (PCR) amplicons (approximately 2.5 kbp) encoding a *cdt* gene operon and two partial and putative open reading frames (ORFs) were identified in six urease-negative (UN) *Campylobacter lari* isolates using a new PCR primer pair constructed *in silico.* Three closely spaced and putative ORFs for *cdtA, cdtB* and *cdtC,* two putative promoters and a hypothetically intrinsic ρ-independent transcription terminator were found in the operon. Each ORF commenced with an ATG start codon and terminated with a TGA stop codon for *cdtA* and *cdtB* and a TAA for *cdtC*. Interestingly, an overlap of four nucleotides was detected between *cdtA* and *cdtB* and the non-coding region of six base pairs occurring between *cdtB* and *cdtC.* The start codons for the three *cdt* genes were preceded by Shine-Dalgarno sequences. Although nucleotide sequence differences were identified at seven loci in the *cdtA* gene, six in *cdtB* and two in *cdtC* among the seven isolates (including *C. lari* RM2100), no polymorphic sites occurred in the putative promoters, hypothetically intrinsic transcription terminator and the three ribosome binding sites among the seven isolates. All nine amino acid residues specific for both *Escherichia coli cdtB* and mammalian DNase I were completely conserved in the *cdtB* gene locus in the 26 *C. lari* isolates, as well as in *C. jejuni* and *C. coli*. No PCR amplicons were generated with urease-positive thermophilic campylobacters (UPTC; $n=10$) using the primer pair.

KEY WORDS: Amino acid sequence.
Base sequence.
Campylobacter lari.
Genes.
Operon.
Polymorphism, genetic.

## Introduction

Thermophilic *Campylobacter* species, primarily *C. jejuni* and *C. coli*, are the recognised causes of acute bacterial diarrhoea worldwide.[1,2] *C. lari* is a relatively recently discovered thermophilic *Campylobacter* species that was first isolated from mammalian and avian species, particularly seagulls of the genus *Larus*.[1,3] *C. lari* has also been shown to be a cause of clinical infection.[4–7] In addition, an atypical urease-positive thermophilic *Campylobacter* (UPTC) group was isolated from the natural environment in England in 1985.[8] Thus, the two representative taxa, namely urease-negative (UN) *C. lari* and UPTC occur within *C. lari.*

Although several *Campylobacter* cytotoxins have been identified,[9,10] only the cytolethal distending toxin (CDT) has been characterised in detail,[11,12] and the *cdt* genes of *C. jejuni* have been cloned and characterised.[11] Cytolethal distending toxin causes progressive cellular distension and ultimately death in several cell lines,[9,11] and also causes cell cycle blockage.[13]

It is suggested that CDTB is a toxic component (features of type I deoxyribonucleases) and that CDTA and C are carriers of CDTB to target cells.[14] Previously, this group has demonstrated genetic heterogeneity of the *cdtB* gene of *C. lari* isolates after polymerase chain reaction (PCR) cloning, sequencing and analysis of the *cdtB* gene fragments (approximately 720 bp in length) from 24 *C. lari* isolates that included eight UPTC isolates.[15] However, in relation to *cdt* genes from *C. lari*, no reports of a full-length *cdt* gene have appeared.

The aim of the present study is to clone, sequence and analyse full-length *cdt* genes, including the putative promoter and terminator structures from UN *C. lari* isolates, and compare their molecular characteristics with those of other thermophilic campylobacters, in order to clarify the full-length *cdt* gene operon in *C. lari* organisms.

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

## Materials and methods

Six isolates of UN *C. lari* analysed in the present study were obtained from different sources (a seagull, humans and mussels) in Japan and Northern Ireland (Table 1).[16] The cells were cultured on blood agar containing defibrinated horse blood (Nippon Bio-Test, Tokyo, Japan) supplemented with a

**Table 1.** UN *C. lari* isolates and other reference strains analysed in the present study, accession numbers of the nucleotide sequence data of the *cdt* gene operon accessible in the DDBJ/EMBL/GenBank and the CMWs of the possible ORFs.

| Isolate no. | *Campylobacter* | Source | Country | Accession number | CMW (Da) | | |
|---|---|---|---|---|---|---|---|
| | | | | | CDTA | CDTB | CDTC |
| JCM2530$^T$ | UN *C. lari* | Seagull | Japan | AB292351 | 30,351 | 29,333 | 21,886 |
| 264 | UN *C. lari* | Mussel | Japan | AB292354 | 30,391 | 29,333 | 21,886 |
| 298 | UN *C. lari* | Human | Japan | AB292355 | 30,391 | 29,333 | 21,886 |
| 84C-1 | UN *C. lari* | Human | N. Ireland | AB292352 | 30,391 | 29,332 | 21,886 |
| 84C-2 | UN *C. lari* | Mussel | N. Ireland | AB292353 | 30,391 | 29,332 | 21,886 |
| 448 | UN *C. lari* | Mussel | N. Ireland | AB292356 | 30,351 | 29,333 | 21,886 |
| RM2100 | UN *C. lari* | Human | USA | NZ_AAFK01000004 | 30,381 | 28,213 | 21,886 |
| RM1221 | *C. jejuni* | Chicken | USA | NC_003912 | 29,946 | 28,942 | 21,157 |
| RM2228 | *C. coli* | Chicken | USA | NZ_AAFL01000003 | 29,205 | 29,006 | 19,919 |

UN *C. lari*: urease-negative *Campylobacter lari*; CMW: calculated molecular weight; JCM: Japan Collection of Microorganisms.

*Campylobacter*-selective medium (Nissui, Tokyo, Japan), under microaerophilic conditions at 37°C for two days.

Template DNA was prepared as described previously[16] after cells were boiled in water for 5 min, and the DNA concentration was adjusted to approximately 300 ng/μL. A schematic representation of the *cdt* genes and their genetic loci for *C. lari* RM2100[17] (DDBJ/EMBL/GenBank Accession No. NZ_AAFK01000004), including the locations of a primer pair for PCR amplification, are shown in Figure 1.

A new PCR primer pair (f-Cl*cdt*all and r-Cl*cdt*all) was designed *in silico* for amplification of the full-length *cdt* gene operon, including the promoter and terminator regions (approximately 2.3 kbp) of *C. lari* based on the sequence information of the *cdt* gene operon and the adjacent genetic loci of *C. lari* RM2100[17] (NZ_AAFK 01000004), as shown in Figure 1.

The PCR mixture contained 10 mmol/L Tris-HCl (pH 8.3), 50 mmol/L KCl, 1.5 mmol/L $MgCl_2$, 400 μmol/L each dNTP, 1 μmol/L each primer, a total of 1 unit of Takara EX Taq DNA polymerase (Takara Bio, Shiga, Japan) and 1 μL template DNA. The PCR method was performed in 50-μL reaction volumes using 30 cycles of 94°C for 1 min, 50°C for 1 min and 72°C for 1.5 min, followed by a final extension at 72°C for 7 min. Amplified PCR products were separated by 1% (w/v) agarose gel electrophoresis in 0.5 x TBE at 100 V and detected by staining with ethidium bromide.

The PCR products amplified by the newly constructed primer pair for the *cdt* gene operon were purified using a QIAquick PCR purification kit (QIAGEN, CA, USA) and inserted in the pGEM-T vector using the pGEM-T Easy Vector System (Promega, Tokyo, Japan). Sequencing of the cloned *cdt* genes was performed using an Hitachi DNA autosequencer (SQ-5500EL) after dideoxynucleotide sequencing using a Thermo Sequenase premixed cycle sequencing kit (Amersham Pharmacia Biotech, Tokyo, Japan). Sequence analysis of the PCR amplicons was carried out using the GENETYX-MAC computer software (Version 9; GENETYX, Tokyo, Japan).

Nucleotide and deduced amino acid sequences of approximately 2.5 kbp regions from the six UN *C. lari* islolates determined in the present study were compared to each other and with the accessible sequence data from some other thermophilic campylobacters using CLUSTAL W software, which was incorporated in the DDBJ.[18]

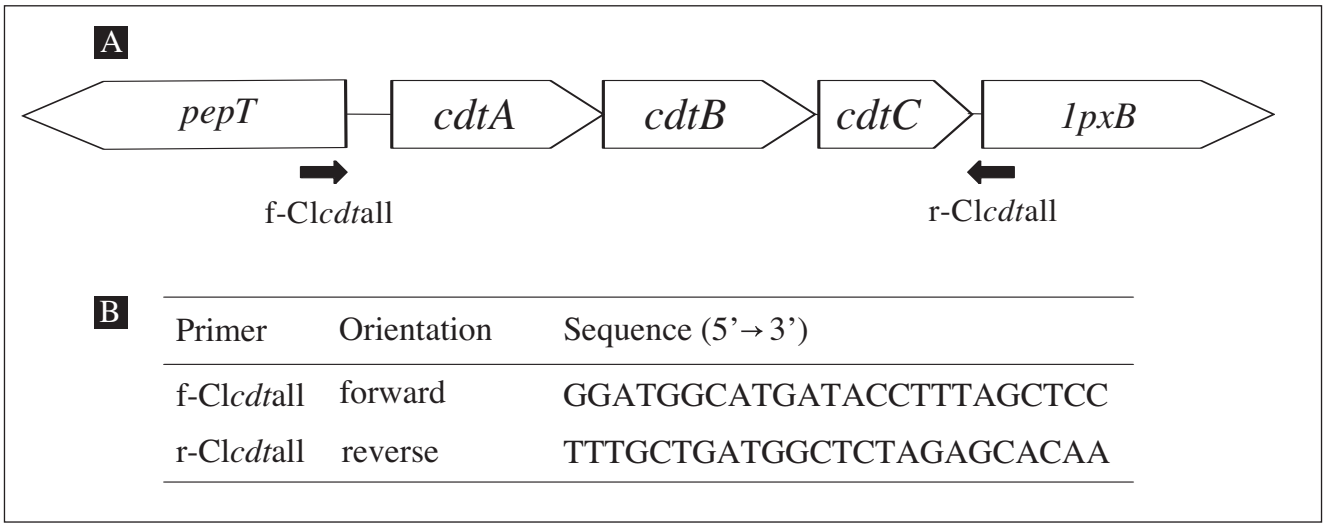


**Fig. 1.** A schematic representation of the *cdt* gene operon and its adjacent genetic loci in *C. lari* RM2100, including the locations of a primer pair for PCR amplification (A) and the primer details (B). *pepT*: peptidase T; *lpxB*: lipid-A-disaccharide synthase.

## Results and discussion

In the present study, the authors first carried out cloning and sequencing of the full-length *cdt* gene operon from six UN *C. lari* isolates using an *in vitro* PCR cloning procedure. The nucleotide sequences (approximately 2.5 kbp in length) determined that contained the full-length *cdt* gene operon from the six isolates are now accessible in the DDBJ/EMBL/GenBank, as shown in Table 1 (AB292351–292356).

Following TA cloning and sequencing, sequences of approximately 2.5 kbp amplicons were identified that encoded a *cdt* gene operon, including two putative promoter structures and a hypothetically ρ-independent intrinsic transcription terminator of approximately 2.3 kbp, and two partial and putative open reading frames (ORFs) immediately upstream and downstream of the operon.

In UN *C. lari* JCM2530$^T$, the present sequence analysis identified three putative ORFs: 810 bp (nucleotide positions [np] 209–1018 bp for *C. lari* JCM2530$^T$) for *cdtA*, 804 bp (np 1018–1821 bp) for *cdtB* and 573 bp (np 1831–2403 bp) for *cdtC*, respectively, all of which were transcribed and translated in the same direction.

The nucleotide positions used were those of UN *C. lari* JCM2530$^T$ (AB292351). Three putative ORFs from the six isolates were similarly predicted to encode peptides of 270, 268 and 191 amino acid residues, with calculated molecular weights (CMW) of approximately 30.4 for CDTA, 29.3 for CDTB and 21.9 kDa for CDTC, respectively (Table 1). Three putative ORFs of *cdtA*, *cdtB* and *cdtC* were identified, based on comparisons of nucleotide and deduced amino acid sequence similarities with those of the corresponding *cdt* genes from *C. jejuni*.[11] Calculated molecular weights of *cdtA, cdtB* and *cdtC* from the seven UN *C. lari* isolates, including *C. lari* RM2100, were shown to be similar to those of *C. jejuni* and *C. coli* strains (Table 1).[11,19]

With regard to the *cdt* genes in *C. lari* JCM2530$^T$, each ORF commenced with an ATG start codon and terminated with a TGA stop codon for *cdtA* and *cdtB*, and a TAA for *cdtC*. The *cdt* gene operon contained a non-coding region of six bp between *cdtB* and *cdtC*. Interestingly, a possible overlap was identified among four nucleotides (ATGA; np 1018–1021 bp), including the three of the stop codon (TGA) of *cdtA* and the start codon (ATG) of *cdtB*. It has also been reported that *cdtA* and *cdtB* genes have a four-nucleotide overlap and the *cdtB* and *cdtC* genes are separated by 10 nucleotides in the *C. jejuni* 81-176 strain.[11]

Probable ribosome-binding (RB) sites (Shine-Dalgarno [SD] sequences)[20,21] that are complementary to a highly conserved sequence of CCUCCU, close to the 3′ end of 16S ribosomal RNA, AGGAGG (np 199–204 bp) for *cdtA*, AGGAG (np 1009–1013 bp) for *cdtB* and AGGAG (np 1823–1827 bp) for *cdtC* and a hypothetical ρ-independent intrinsic transcriptional terminator sequence (np 2413?2428 bp) were identified in *C. lari* JCM2530$^T$ (Fig. 2). The hypothetical intrinsic transcription terminator structure, which contains a G+C-rich region near the base of the stem and a single-stranded run of T residues, was identified immediately downstream of *cdtC* (Fig. 2). The start codons of the three *cdt* genes were preceded by SD sequences.

Two putative promoter structures, consisting of sequences at the -35-like (TTTACT; np 159–164 bp) and -10 (TTATATT; np 184–190 bp) regions, as well as the start codon (ATG; np 209–211 bp), were identified for the transcriptional promoter immediately upstream of the *cdtA* gene in *C. lari* JCM2530$^T$. Thus, the *cdtA*, *cdtB* and *cdtC* genes identified in the present study in the six UN *C. lari* isolates appear to form an operon and to function in the cell.

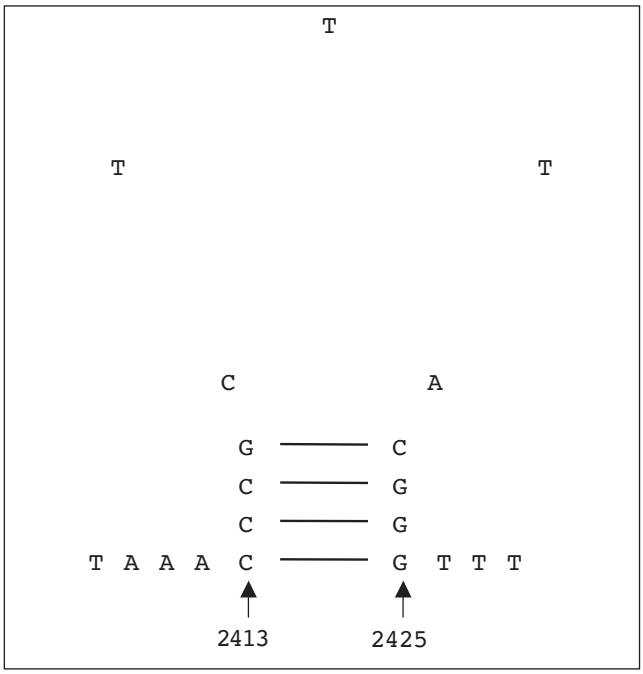


**Fig. 2.** A hypothetically intrinsic ρ-independent transcription terminator structure which contains a G+C-rich region near the base of the stem (2413–2425 bp) and a single-stranded run of T residues (2426–2428 bp) for UN *C. lari* JCM 2530$^T$.

Nucleotide sequence differences were identified at seven loci in the *cdtA* gene (np 278, 461, 515, 644, 671, 818 and 865 bp) and at six positions (np 1212, 1233, 1476, 1659, 1732 and 1803 bp) in the *cdtB* gene among the seven UN *C. lari* isolates, including the *C. lari* RM2100 strain examined. Nucleotide sequence differences were also demonstrated at two loci in *cdtC* (np 1987 and 2325 bp). However, no polymorphic sites were identified in the putative promoter, the hypothetical intrinsic transcription terminator regions and three RB sites among the seven isolates.

Thus, regarding the *cdt* gene operon, three *cdt* structural genes (*cdtA, cdtB* and *cdtC*) carried some polymorphic sites, but no polymorphic sites were demonstrated in the regions, promoter, terminator and RB sites in the seven UN *C. lari* isolates. In addition, the putative ORFs for *cdtA* from seven UN *C. lari* isolates showed 98.5–100% amino acid sequence similarities, with 95.9–100% for *cdtB* and 100% for *cdtC* (data not shown). The putative ORFs for *cdtA, cdtB* and *cdtC* from the seven isolates showed 50.5–66.5% amino acid sequence similarities to that of the *C. jejuni* RM1221 strain (NC_003912) (data not shown).

In relation to the adjacent genetic loci of the *cdt* gene operon analysed, a partial and putative ORF (np 1–63 bp) of the peptidase T (*pepT*) gene and two putative promoter structures (-35 region [np 128–133 bp] and -10 region [np 92–97 bp]) and a RB site (TCCT; np 70–73 bp) were shown to occur in the reverse direction immediately upstream of the *cdt* gene operon, in all seven isolates of UN *C. lari* examined. In addition, another partial and putative ORF (np 2485–2501 bp) of the lipid-A-disaccharide synthase (*lpxB*) gene and two

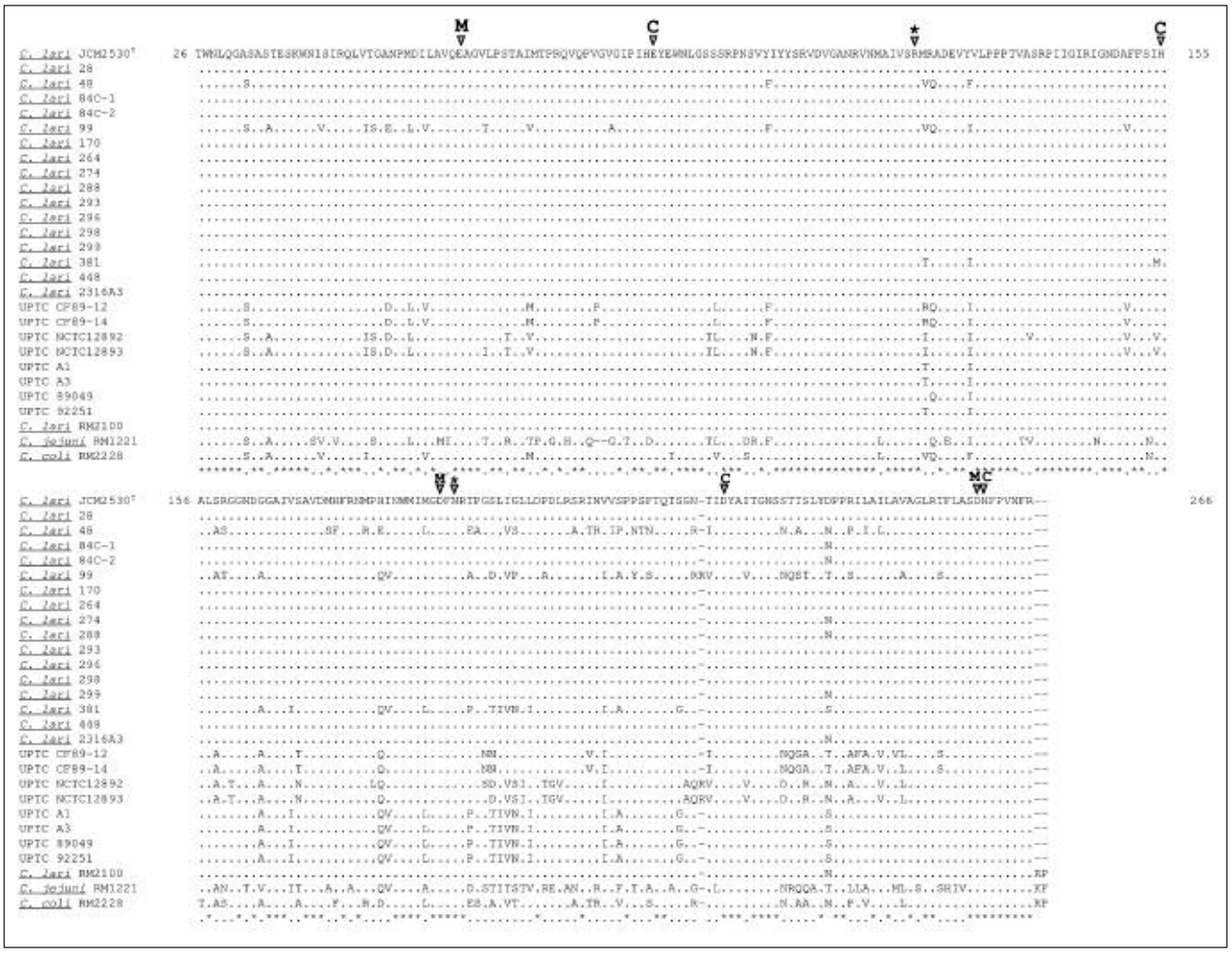

**Fig. 3.** Amino acid sequence alignment analysis of the putative ORF for the *cdtB* gene from the 25 *C. lari* isolates. Some of the sequences used in this alignment are derived from a previous report.[15] Amino acid sequences of those from the three *Campylobacter* reference strains are aligned for comparison. M: metal ion-binding residues; C: catalytic residues; *DNA contact residues.
Numbers at the left and right refer to the amino acid positions in the putative ORF for *cdtB* of *C. lari* JCM2530$^T$ (AB292351).

putative promoters (-35 region of TTTAAA [np 2426–2431 bp] and -10 region of TAAAAT [np 2462–2467 bp]) and an RB site (GTAGG; np 2477–2481 bp) were also shown to occur immediately downstream of the *cdt* gene operon.

Regarding the putative ORF for *cdtB* gene (804 bp), the authors have already cloned, sequenced and analysed the *cdtB* gene fragments of approximately 720 bp from 24 isolates (UN *C. lari* [$n$=16], UPTC [$n$=8]) obtained from different sources and in various countries,[15] and the genetic heterogeneity of the *cdtB* gene locus among the isolates of *C. lari* was demonstrated. The nucleotide sequences of the partial *cdtB* gene fragments of 16 UN *C. lari* isolates showed 81–100% similarities.[15] In five UN *C. lari* isolates (JCM2530$^T$, 264, 298, 84C-1 and 84C-2) examined in the present study, identical polymorphic sites were confirmed to occur within the approximate 720 bp segment.

In addition, Elwell and Dreyfus reported that *cdtB* of *Escherichia coli* shows significant position-specific homology to type I mammalian DNases at specific DNase I residues involved in enzyme catalysis, DNA binding and metal ion binding.[22] However, in the present study, all nine amino acid residues were completely conserved in the *cdtB* gene locus in the 26 *C. lari* isolates examined, including in *C. lari* RM2100, *C. jejuni* and *C. coli* (Fig. 3).

In the present study, the authors attempted to amplify the corresponding segment using the primer pair with the genomic DNAs from the isolates of UPTC[23] as templates. However, no PCR amplicons could be generated from at least 10 isolates (data not shown). This may indicate that the primer (f-Cl*cdt*all and r-Cl*cdt*all) target regions of the UPTC isolates are not completely conserved, preventing perfect hybridisation to the primers used. Thus, new approaches are required to analyse full-length *cdt* genes from UPTC organisms.

Consequently, the previous[15] and present results clearly suggest that *C. lari* isolates (UN *C. lari* and UPTC) commonly carry a *cdt* gene operon in their genomes. □

*The research was supported in part by The Promotion and Mutual Aid Corporation for Private Schools of Japan, and by Grant-in-Aid for Matching Fund Subsidy for Private Universities. MM and JEM were funded through a Great Britain Sasakawa Foundation Award to examine the clinical significance of* Campylobacter *infection in the UK and Japan.*

## References


1 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.

2 Blaser MJ, Taylor DN, Feldman RA. Epidemiology of *Campylobacter jejuni* infections. *Epidemiol Rev* 1983; **5**: 157–76.

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

4 Nachamkin I, Stowell C, Skalina D, Jones AM, Hoop RM 2nd, Smibert RM. *Campylobacter laridis* causing bacteremia in an immunosuppressed patient. *Ann Int Med* 1984; **101**: 55–7.

5 Simor AE, Wilcox L. Enteritis associated with *Campylobacter laridis. J Clin Microbiol* 1987; **25**: 10–2.

6 Martinot M, Jaulhac B, Moog R *et al*. *Campylobacter lari* bacteremia. *Clin Microbiol Infect* 2001; **7**: 96–7.

7 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.

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

9 Johnson WM, Lior H. A new heat-labile cytolethal distending toxin (CLDT) produced by *Campylobacter* spp. *Microb Pathog* 1988; **4**: 115–26.

10 Schulze F, Hanel I, Borrmann E. Formation of cytotoxins by enteric *Campylobacter* in humans and animals. *Zentralbl Bakteriol* 1998; **288**: 225–36.

11 Pickett CL, Pesci EC, Cottle DL, Russell G, Erdem AN, Zeytin H. Prevalence of cytolethal distending toxin production in *Campylobacter jejuni* and relatedness of *Campylobacter* spp. *cdtB* genes. *Infect Immun* 1996; **64**: 2070–8.

12 Pickett CL, Whitehouse CA. The cytolethal distending toxin family. *Trends Microbiol* 1999; **7**: 292–7.

13 Whitehouse CA, Balbo PB, Pesci EC, Cottle DL, Mirabito PM, Pickett CL. *Campylobacter jejuni* cytolethal distending toxin causes a G2 phase cell cycle block. *Infect Immun* 1998; **66**: 1934–40.

14 Lara-Tejero M, Galan JE. A bacterial toxin that controls cell cycle progression as a deoxyribonuclease 1-like protein. *Science* 2000; **290**: 354–7.

15 Shigematsu M, Harada Y, Sekizuka T *et al*. Genetic heterogeneity of the cytolethal distending toxin (*cdt*)*B* gene locus among isolates of *Campylobacter lari. Br J Biomed Sci* 2006; **63**: 179-81.

16 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* 2004; **155**: 185–91.

17 Fouts DE, Mongodin EF, Mandrell RE *et al*. Major structural differences and novel potential virulence mechanisms from the genomes of multiple *Campylobacter* species. *PLoS Bio*l 2005; **3** (1): e15 (Epub 4 Jan 2005).

18 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 Acid Res* 1994; **22**: 4673–80.

19 Eyigor A, Dawson KA, Lasnglois BE, Pickett CL. Detection of cytolethal distending toxin activity and *cdt* genes in *Campylobacter* spp. isolated from chicken carcasses. *App Environ Microbiol* 1999; **65**: 1501–5.

20 Shine J, Dalgarno L. The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc Natl Acad Sci USA* 1974; **71**: 1342–6.

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

22 Elwell CA, Dreyfus LA. DNase 1 homologous residues in CdtB are critical for cytolethal distending toxin-mediated cell cycle arrest. *Mol Microbiol* 2000; **37**: 952–63.

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