Molecular structural analysis of major outer membrane protein (MOMP) gene clusters in *Campylobacter lari*

T. NAKAJIMA[*], W. ARA[*], K. SAITO*, J. E. MOORE[†§], B. C. MILLAR[†] and M. MATSUDA[*]

[*]*Laboratory of Molecular Biology, Graduate School of Environmental Health Sciences, Azabu University, Japan;* [†]*Department of Bacteriology, Northern Ireland Public Health Laboratory, Belfast City Hospital;* [‡]*School of Biomedical Sciences, University of Ulster, Cromore Road, Coleraine; and* [§]*Centre for Infection and Immunity, Queen's University, Belfast, Northern Ireland, UK*

Accepted: 3 January 2014

ABSTRACT

Southern hybridisation shows that urease-negative (UN) *Campylobacter lari* JCM2530[T] carries two putative major outer membrane protein (MOMP) genes. Sequences of approximately 2.1 kbp, encoding non-coding (NC) regions, with possible open reading frames (ORFs) for MOMP (*porA1* or *porA2*) of approximately 1.2 kbp, NC regions and partial and putative Cla_0435 or Cla_1109 ORFs were identified in all five UN *C. lari* isolates examined, following polymerase chain reaction (PCR) cloning and sequencing. Each putative MOMP structural gene carried start and stop codons and ribosome binding sites of 1236–1278 bp in length. The putative σ^{70} transcriptional promoter and the hypothetical ρ-independent transcription terminator structures were also seen. Using Northern hybridisation, there was *in vivo* monocistronic MOMP gene transcription. In addition, in a Japanese urease-positive thermophilic *Campylobacter* (UPTC) CF89-12 strain, the *porA1* gene locus, including an extra gene (approximately 2000 bp in length) was identified. The extra gene may occur within the *porA1* gene locus in the eight UPTC isolates of the 23 *C. lari* isolates examined. Thus, a genetic heterogeneity occurred within the *porA1* gene locus from some of the *C. lari* organisms including the UPTC CF89-12.

KEY WORDS: Campylobacter lari.
Major outer membrane protein genes.
Northern blot hybridisation.
Southern blot hybridisation.

Introduction

Thermophilic *Campylobacter* species, primarily *Campyobacter jejuni* and *C. coli*, are major recognised causes of acute bacterial diarrhoea.[1,2] In addition, *C. lari* is a thermophilic *Campylobacter* species that was first isolated from mammalian and avian species, particularly seagulls of the genus *Larus*.[2,3] *C. lari* has also been shown to be a cause of clinical infection, although cases are less common than either *C. jejuni* or *C. coli*.[4–6]

Moreover, an atypical urease-producing group of isolates, namely urease-positive thermophilic *Campylobacter* (UPTC), has been isolated from the natural environment in England in 1985.[7] Thereafter, UPTC organisms were reported as a biovar or variant of *C. lari*.[8,9] Additional isolates of UPTC have also been reported in several countries.[10–13] Thus, these two representative taxa, namely urease-negative (UN) *C. lari* and UPTC, occur within the *C. lari* species.[14]

Porins are integral outer membrane proteins of pore formation of a hydrophilic nature in Gram-negative bacteria and are involved in adaptation of bacteria to their environment.[15] In *C. jejuni*, to date, two porins have been characterised, named the major outer membrane protein (MOMP)[16] and Omp50 (a minor porin).[17] The former is a large-channel porin,[18] encoded by the *porA* gene.[19] Labesse *et al*. compared seven MOMP sequences (*C. jejuni* [*n*=5], *C. coli* [*n*=2]) following cloning and primary structural characterisation and showed an alteration of preserved and divergent regions.[20] In addition, more than 100 MOMP sequences from *C. jejuni* and *C. coli* have been deposited in the DDBJ/EMBL/GenBank databases (data not shown). Although two MOMP sequences of *porA1* (Cla_0434) and *porA*2 (Cla_1108) have been identified in the human clinical UN *C. lari* RM2100 strain, following whole-genome shotgun sequencing (AAFK01000002)[21] and complete genome sequencing (CP000932)[22] analyses, as yet no reports have been published.

This study aims to clone, sequence and molecularly characterise full-length gene sequences encoding the MOMP, including the promoter and terminator regions from UN *C. lari* isolates and to compare the sequences with those from the other thermophilic campylobacters, in order to clarify MOMP structure in *C. lari* organisms. Subsequently, it aims to clarify whether or not the expression of MOMP genes occurs in the UN *C. lari* cells using Northern blot hybridisation, and attempt to molecularly identify and characterise the *porA1* and *porA*2 gene loci within UPTC organisms.

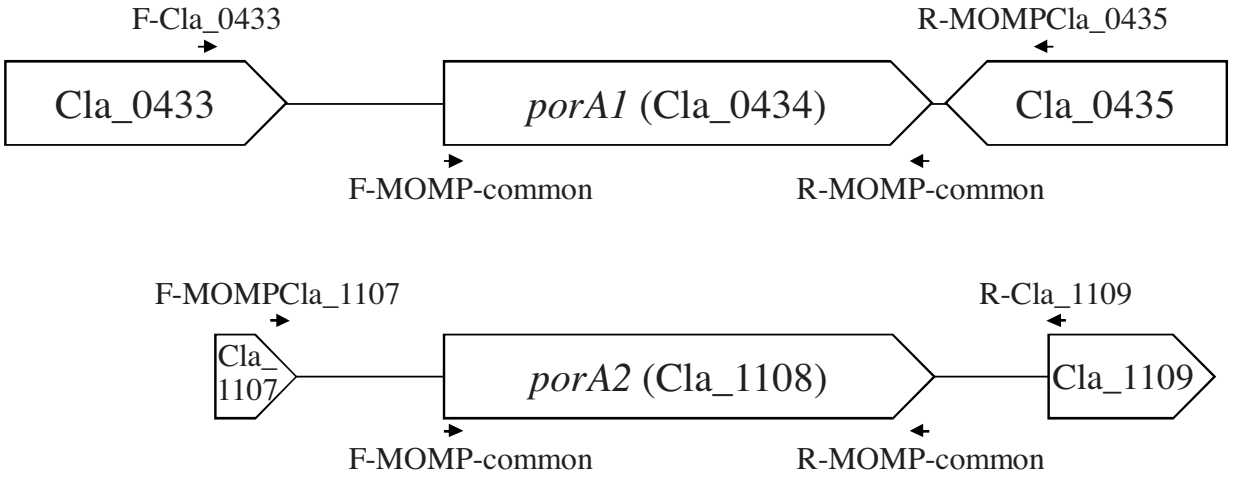
Materials and methods

Isolates and culture condition

A total of 24 *C. lari* isolates (UN *C. lari* [*n*=14], UPTC [*n*=10]) obtained from various sources and in several countries in

Correspondence to: Professor Motoo Matsuda
Laboratory of Molecular Biology, School of Environmental Health Sciences
Azabu University, 252-5201, Japan
Email: matsuda@azabu-u.ac.jp

A



B

Primer	Nucleotide sequence (5'-3')
F-MOMP-common	ATGAAACTAGTTAAACTTAGTTTAG
R-MOMP-common	TTAGAATTTGTAWAGAGCTTSAAGTC
F-Cla_0433	GGTTTAGTTTTTGATGATTATGGTCG
R-MOMPCla_0435	TTAATAATCGATCTTAGAGC
F-MOMPCla_1107	GCTATAGAAGCTTTAGAGGATC
R-Cla_1109	GATTTTACATTAAAAGCYAATAG

Fig. 1. Schematic representations of putative full-length *porA1* and *porA2* gene arrangements and their adjacent genetic loci in *C. lari* RM2100, including the locations of the primer pairs used for PCR amplifications (A) and primer sequences employed (B).

Asia, Europe and North America were analysed (Table 1). These isolates were cultured on Mueller-Hinton broth medium 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 using BBL Campypak Microaerophilic System envelopes (Becton Dickinson, NJ, USA).

Genomic DNA preparation

Genomic DNA was prepared using sodium dodecyl sulphate, proteinase K treatment, phenol-chloroform extraction, and ethanol precipitation.[23]

Southern blot hybridisation analysis

Southern blot hybridisation analysis for the MOMP gene(s) was carried out using digoxigenin (DIG)-labelled MOMP structural gene fragment (approximately 960 bp) prepared from UN *C. lari* JCM2530^T strain as a probe with *Pst I*-digested whole genomic DNA, according to the procedure described by Sambrook and Russell.[24] The fragment probe was amplified using a degenerate primer pair (F-/R-MOMP-common), designed *in silico* (Fig. 1). Random primer extension was performed in order to prepare the fragment probe using the Random Primer DNA Labelling Kit (Version. 2) (TaKaRa Bio, Shiga, Japan.).

Primer design and PCR amplification

Two primer pairs (F-Cla_0433/R-MOMPCla_0435 for *porA1*; F-MOMPCla_1107/R-Cla_1109 for *porA2*) were prepared *in silico* for amplification of an approximately 2.1 kbp sequence including the full-length MOMP gene (approximately 1.2 kbp) and its adjacent genetic loci, based on the sequence information of the *C. lari* RM2100 strain (Fig. 1).[22] Schematic representation of the MOMP genes and their adjacent genetic loci for UN *C. lari*, including the locations of the primer pairs and nucleotide sequences of those constructed *in silico* for polymerase chain reaction (PCR) amplification, are shown in Figure 1.

The PCR mixture contained 10 mmol/L Tris-HCl (pH 8.3), 50 mmol/L KCl, 3.0 mmol/L $MgCl_2$, 400 µmol/L each dNTP, 0.5 µmol/L each primer, 50 ng template DNA and a total of 1.25 units rTaq DNA polymerase (TaKaRa Bio). The PCR was performed in 50 µL reaction volumes at 94°C for 2 min, 30 cycles of 94°C for 30 sec, 50°C for 30 sec and 72°C for 150 sec, followed by a final extended elongation step at 72°C for 3 min.

Product purification and cycle sequencing

The PCR products separated by 1% (w/v) agarose gel electrophoresis in 0.5× TBE were purified using ExoSAP-IT (GE Healthcare, Tokyo, Japan) at 37°C for 30 min, and then inactivated at 80°C for 15 min. The purified fractions were subjected to cycle sequencing with BigDye Terminator

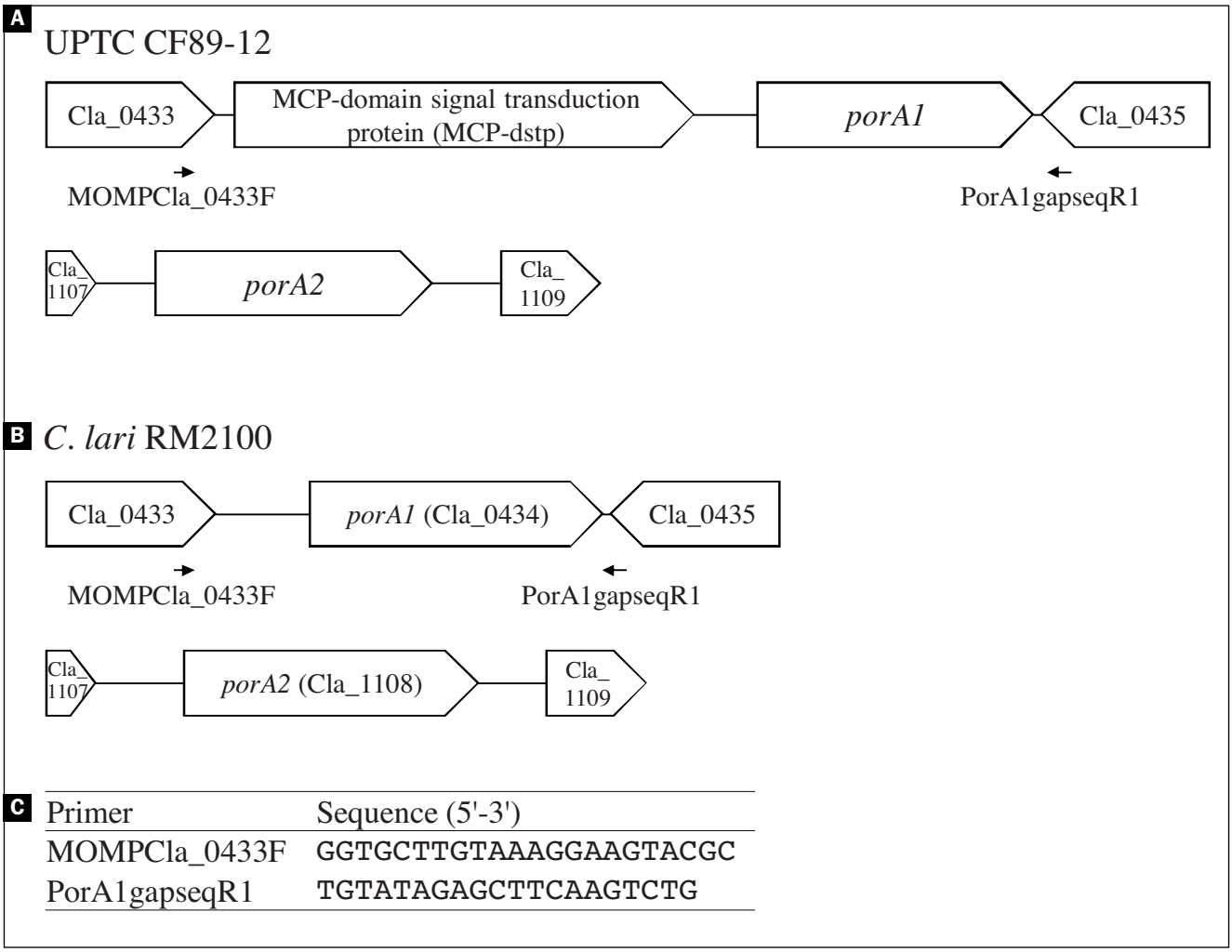


Fig. 2. Schematic representations of the putative *porA1* and *porA2* structural genes and their adjacent genetic loci in the UPTC CF89-12 (A) and UN *C. lari* RM2100 (B), including the locations of the primer pair for PCR amplification of the *porA1* structural gene locus (A and B) and primer sequences employed (C).

(Applied Biosystems, Tokyo, Japan) and with the sequencing primers. The reaction products were separated and detected on an ABI PRISM 3100 genetic analyser (Applied Biosystems).

The nucleotide sequences of the MOMP structural genes from the five UN *C. lari* isolates (JCM2530^T, 299, 300, 448, 84C-1) were analysed using the GENETYX computer software (Version 9; GENETYX, Tokyo, Japan). Nucleotide sequences of the MOMP structural genes and their possible open reading frames (ORFs) from the UN *C. lari* isolates were compared to each other and to accessible sequences data employing CLUSTAL W software (1.7 Program),[25] which was incorporated in the DDBJ.

Total RNA extraction, purification and Northern blot hybridisation

Total cellular RNA was extracted and purified from the UN *C. lari* JCM2530^T strain cells using RNAprotect Bacteria Reagent and RNeasy Mini Kit (Qiagen, Tokyo, Japan). Northern blot hybridisation analysis was carried out according to the procedure described by Sambrook and Russell[24] using a PCR-amplified, DIG-labelled MOMP gene fragment from the *C. lari* JCM2530^T as a probe. The fragments were amplified using a primer pair of F-/R-MOMP-common, and a random primer extension procedure was employed for the preparation of the probe, as described above.

Construction of genomic DNA library, nucleotide sequencing and sequence analysis

Construction of the genomic DNA library of UPTC CF89-12 strain cells, nucleotide sequencing and sequence analysis were carried out using the procedures described previously.[26]

Amplification of the* porA1 *locus

A primer pair of MOMPCla_0433F and PorA1 gapseqR1 (Fig. 2) was designed in order to amplify the *porA1* locus between the Cla_0433 and the *porA1* in the *C. lari* isolates, based on the nucleotide sequence information of the regions from the *C. lari* isolates examined (Table 1).

Results

Southern blot hybridisation

The study attempted to examine the possible numbers of the putative *porA* gene in UN *C. lari* isolate genomic DNA using

Table 1. *C. lari* isolates and other thermophilic *Campylobacter* reference strains analysed in the present study and their accession numbers of the nucleotide sequence data accessible in the DDBJ/EMBL/GenBank.

Campylobacter isolate	Source	Country	Accession number
UN *C. lari* JCM2530^T	Seagull	Japan	AB569412 AB569413
UN *C. lari* 28	Mussel	N Ireland	NA
UN *C. lari* 170	Seagull	Japan	NA
UN *C. lari* 175	Seagull	Japan	NA
UN *C. lari* 176	Black-tail gull	Japan	NA
UN *C. lari* 264	Mussel	N Ireland	NA
UN *C. lari* 274	Mussel	N Ireland	NA
UN *C. lari* 295	Human	NA	NA
UN *C. lari* 298	Human	Canada	NA
UN *C. lari* 299	Human	USA	AB569418 AB569419
UN *C. lari* 300	Seagull	USA	AB569420 AB569421
UN *C. lari* 382	Mussel	N Ireland	NA
UN *C. lari* 448	Mussel	N Ireland	AB569416 AB569417
UN *C. lari* 84C-1	Human	N Ireland	AB569414 AB569415
UPTC CF89-12	River water	Japan	AB736316 AB736317
UPTC A1	Seagull	N Ireland	NA
UPTC 89049	Human	France	NA
UPTC 2	Oyster	N Ireland	NA
UPTC 27	Mussel	N Ireland	NA
UPTC 136	Scallop	N Ireland	NA
UPTC 150	Cockle	N Ireland	NA
UPTC 182	Sea water	N Ireland	NA
UPTC 476	Mussel	N Ireland	NA
UPTC 504	Mussel	N Ireland	NA
C. lari RM2100	Human	USA	AAFK01000002
C. jejuni NCTC11168	Human	USA	NC_002163
C. jejuni RM1221	Chicken	USA	NC_003912
C. coli RM2228	Chicken	USA	AAFL01000008

NA: not available

Southern blot hybridisation. As shown in Figure 3, *PstI*-digested whole genomic DNA prepared from the UN *C. lari* JCM2530^T gave two positive hybridisation signals, suggesting that UN *C. lari* isolates carry two putative MOMP genes or homologues (*porA1* and *porA2*) within the genomic DNA.

Sequencing and analysis

As Southern blot hybridisation analysis suggested two putative MOMP genes or homologues in the UN *C. lari* isolate cells, an attempt was made to sequence and analyse two putative MOMP genes from the five isolates (UN *C. lari* JCM2530^T, 299, 300, 448 and 84C-1; Table 1), and to amplify the *porA1* gene and its adjacent genetic loci using the F-Cla_0433/R-MOMPCla_0435 primer pair, and the *porA2* gene and its adjacent genetic loci using the F-MOMPCla_1107/R-Cla_1109 primer pair (Fig. 1). The PCR products amplified for *porA1* and *porA2* with the five UN *C. lari* isolates (Fig. 4) were purified and the nucleotide and deduced amino acids sequences were determined. These are accessible in the DDBJ/EMBL/GenBank nucleotide sequences database (Table 1).

The sequences of approximately 2.1 kbp encoding the approximately 400 bp non-coding (NC) region, the possible ORFs of the two putative MOMP genes (*porA1* and *porA2*) of an approximate 1.2 kbp, approximately 20 bp or 290 bp NC regions and a Cla_0435 ORF of approximately 740 bp or a Cla_1109 ORF of approximately 430 bp, were identified within the *porA1* and *porA2* loci in the six UN *C. lari* isolates including the *C. lari* RM2100 strain. Possible *porA1* and *porA2* ORFs and their calculated molecular weights (CMWs) from the six UN *C. lari* isolates including *C. lari* RM2100 examined are summarised in Table 2.

Nucleotide sequence analysis

Two NC regions of approximately 400 bp for *porA1* and *porA2* and 20 bp for *porA1* and 290 bp for *porA2* occurred upstream and downstream of the MOMP gene ORFs, respectively, within the UN *C. lari* RM2100 strain. Nucleotide sequences of approximately 200 bp of the NC regions upstream of the ORFs from the five UN *C. lari* isolates and *C. lari* RM2100 showed 96.7–100% sequence similarity (Table 4). In these regions, typical promoter consensus sequences at the –10-like region (TATAAT) were shown at the locus (between

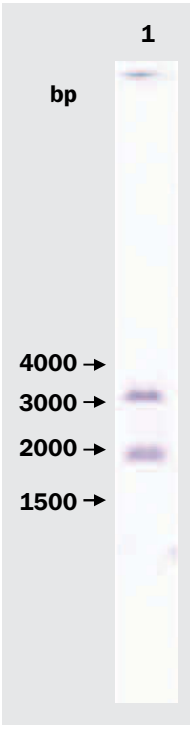


Fig. 3 (Left). Southern blot hybridisation analysis of the whole genomic DNAs digested with Pst I from the UN *C. lari* JCM2530^T using MOMP gene fragment as a probe (Lane 1).

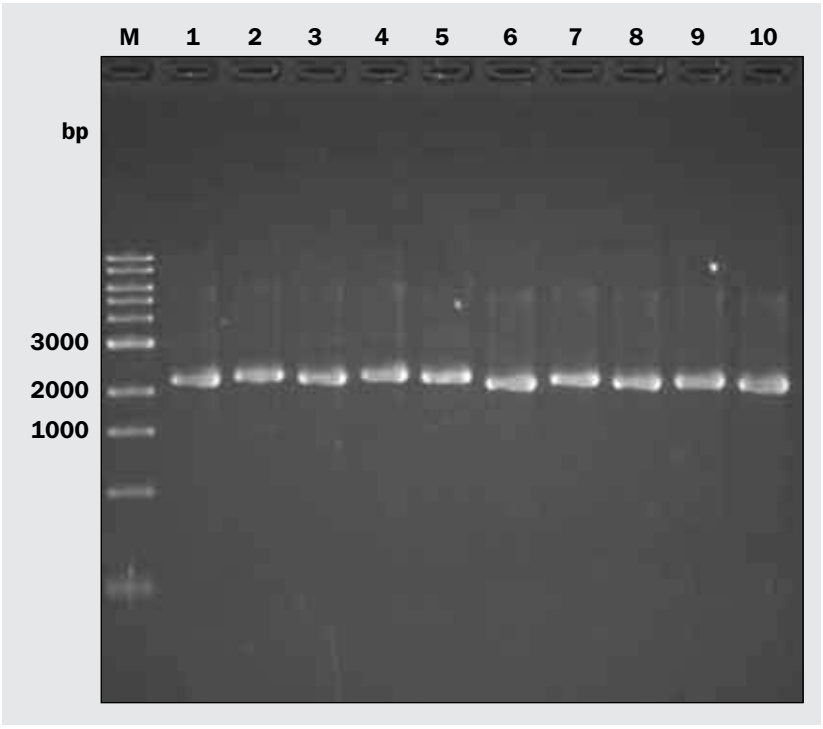


Fig. 4 (Right). Agarose gel electrophoresis profiles of PCR products amplified using the two primer pairs shown in Figure 1, with five UN *C. lari* isolates. Lanes 1–5: employing a primer pair of F-Cla_0433/R-MOMPCla_0435; lanes 6–10, F-MOMPCla_1107/ R-Cla_1109. Lane M, 1 kbp DNA ladder (New England BioLabs, Tokyo, Japan); Lanes 1 and 6: *C. lari* JCM2530^T; Lanes 2 and 7: 84C-1; Lanes 3 and 8: 448; Lanes 4 and 9: 299; Lanes 5 and 10: 300.

nucleotide positions [np] 598–603 bp for the *porA1* of the UN *C. lari* JCM2530^T) in all six isolates of UN *C. lari* (Fig. 5). The nucleotide positions used are for those of the *porA1* or *porA2* from the UN *C. lari* JCM2530^T, and the others are indicated.

No consensus sequence at the –35-like region was identified, and a semi-conserved T-rich region was identified (between np 560 and 573 bp for the *porA1*; T, 11/14) for both *porA1* and *porA2* (Fig. 5), instead of the region shown in *RpoD* promoters in the genome of *C. jejuni*.[27] Therefore, these two MOMP genes may be transcribed by the σ^{70} factor in the UN *C. lari* organisms. Thus, the putative σ^{70} transcriptional promoter structures were identified for the MOMP structural genes within all six UN *C. lari* isolates examined.

Probable ribosome-binding (RB) sites[28] that are complementary to the highly conserved sequences of CCUCCU, close to the 3′ end of 16S rRNA AGGAG, were also identified (between np 629 and 633 bp for *porA1*) in all six isolates (Fig. 5).

In relation to the transcription terminator for MOMP genes, hypothetical ρ-independent transcription terminator structures were identified (between np 2049 and 2073 bp for

```
C. lari JCM2530T porA1      426 GAATAAAATTA-AATTATAATTTTAAAAATTTA-ATATAATATAGTTTGATTATCTA-AACTATAATCATTATATA-CTAATAATTTAT-AGTACAACAACATACTATAACTTCAAATAT
C. lari JCM2530T porA2      242 ..........-..........-..........-..........-..........-..........-..........-..........-..........
C. lari 299 porA1           426 ..........-.....G....-..........-..........-..G.....-..........-..........-..........-..........
C. lari 299 porA2           242 ..........-.....G....-..........-..........-..........-..........-..........-..........-..........
C. lari 300 porA1           427 ..........-..........-..........-..........-..........-..........-..........-.....C....-..........
C. lari 300 porA2           243 ..........-..........-..........-..........-..........-..........-..........-.....C....-..........
C. lari 448 porA1           426 ..........-.....G....-..........-..........-..........-..........-..........-.....C....-.....C.C
C. lari 448 porA2           242 ..........-.....G....-..........-..........-..........-..........-..........-.....C....-.....C.C
C. lari 84c-1 porA1         426 ..........-..........-..........-..........-..........-..........-..........-..........-..........
C. lari 84c-1 porA2         242 ..........-..........-..........-..........-..........-..........-..........-..........-..........
C. lari RM2100 porA1    409,952 ..........-.....G....-..........-..........-..........-..........-..........-..........-..........
C. lari RM2100 porA2  1,059,128 ..........-.....G....-..........-..........-..........-..........-..........-..........-..........
UPTC CF89-12 porA1        2,862 ..GC.T.....C.T..G...AAGC...C...C.T.G....-C.CC..A.T..T.AA.G.T...GTT...AAC..G.T..TTA..GA...GTC..T...T..T..-...T.A.A.-T....
UPTC CF89-12 porA2          254 ..GC.T.....C.T..G...AAGC...C...C.T.G....-C.CC..A.T..T.AA.G.T...GTT...AAC..G.T..TTA..GA...GTC..T...T..T..-...T.A.A.-T....
                                **  * ***** * ** ***    *** *** * * ****  *  ** * ** *  * * ***   ***   ** * **   **  ***   ** *** **  * *** * * *  * *

C. lari JCM2530T porA1      541 ATCCCAATAAATACATATA---TTTTTATTTACTTTAAGAATTCTCATATAAAACTATGGTATAATACAAGCCATAGTTTTAATTTTAATAAGGAGAAATCTCATGAAACTAGTTAAACT
C. lari JCM2530T porA2      357 ..................---.....................................................................T.....................
C. lari 299 porA1           541 .............--...---.....G.....................A.....................................T.....................
C. lari 299 porA2           357 .............--...---.....G.....................A.....................................T.....................
C. lari 300 porA1           542 .............--...---.....G.....................................................T.....................
C. lari 300 porA2           358 .............--...---.....G.....................................................T.....................
C. lari 448 porA1           541 ..T..............---.....G.....................................................T.....................
C. lari 448 porA2           357 ..T..............---.....G.....................................................T.....................
C. lari 84c-1 porA1         541 ..................---.....................................................T.....................
C. lari 84c-1 porA2         357 ..................---.....................................................T.....................
C. lari RM2100 porA1    410,067 .............--...---.....G.....................A.....................................T.....................
C. lari RM2100 porA2  1,059,243 .............--...---.....G.....................A.....................................T.....................
UPTC CF89-12 porA1        2,979 .AT.T....T..TTT....TTG.....................T.....C.....................................T.....................
UPTC CF89-12 porA2          371 .AT.T....T..TTT....TTG.....................T.....C.....................................T.....................
                                *  * **** **    ***   ***** ********** ********** ********** ********** ********** ********** **********
```

Fig. 5. Nucleotide sequence alignment analysis of the non-coding regions upstream of *porA1* and *A2* structural genes including the putative promoter structures, consisting of a semi-conserved T-rich and a -10 like regions, as well as the start codon ATG. Promoter structures, RB site and start codon in the UN *C. lari* JCM2530^T *porA1* are underlined. Dots indicate identical bases; changes are indicated; dashes are deleted bases; identical positions in all cases are marked by asterisks; numbers at the left and right refer to the nucleotide positions in each isolate.

A

```
C. lari JCM2530T porA1        2,001   CATGTATATTTTTACTTGCTAGTATTCTTAATATTTTACCCCTATAAGCCTTTGCAAAGTGACTTAAGGTTTTGGAATTTTTTAAAAAAGTGAGGGTAA
C. lari JCM2530T porA2        1,817   ....................................................................................................
C. lari 299 porA1             2,041   ....................................................................................................
C. lari 299 porA2             1,857   ....................................................................................................
C. lari 300 porA1             2,003   ..........................................................G.........................................
C. lari 300 porA2             1,819   ..........................................................G.........................................
C. lari 448 porA1             2,043   ....................................................................................................C..
C. lari 448 porA2             1,859   ....................................................................................................
C. lari 84c-1 porA1           2,001   ....................................................................................................
C. lari 84c-1 porA2           1,817   ....................................................................................................
C. lari RM2100 porA1        411,567   T...................A..C.....G.........................A........A.........................A....
C. lari RM2100 porA2      1,060,743   T...................A..C.....G.........................A........A.........................A...G
UPTC CF89-12 porA1            4,412   ....................................................................................................
UPTC CF89-12 porA2            1,804   ....................................................................................................
                                      ******************** ** ***** ********************** ** ******** ************************ * *
```

porA1) for both the *porA1* and *porA2* genes in all six UN *C. lari* isolates (Fig. 6A). The hypothetically intrinsic transcription terminator structure that contains a G+C-rich region near the base of the stem and a single-stranded run of U residues (between np 2049 and 2073 bp for *porA1*) were identified for the two MOMP genes within at least the five UN *C. lari* isolates examined in the present study (Fig. 6A and B). In addition, the *porA1* and *porA2* gene loci are located approximately 649 kbp away from each other within the UN *C. lari* RM2100 genome DNA (Figs. 5 and 6A).

Northern blot hybridisation analysis

Northern blot hybridisation analysis detected MOMP gene transcription in the UN *C. lari* JCM2530^T (Fig. 7). As the positive signal of the hybridisation was shown at around 1.4 kbp, the MOMP genes could be transcribed monocistronicly. Thus, *in vivo* MOMP gene transcription may occur in UN *C. lari* organisms.

Nucleotide sequence of the* porA1 *gene locus

During genome sequence analysis for a representative taxon of *C. lari* UPTC CF89-12, the *porA1* (a structural gene 1206 bp in length; np 3082–4287 bp; AB736316) gene locus was demonstrated. Therefore, the sequence following the gap closing (approximately 4000 bp in length) was analysed. The *porA1* locus of the UPTC CF89-12 analysed in the present study is schematically illustrated in Figure 2A. As shown, the *porA1* gene locus containing the putative *MCP-dstp* gene (2001 bp in length) occurring upstream of the gene is different from that in UN *C. lari* RM2100 (Fig. 2). The possible ORFs of this putative *MCP-dstp* gene in the UPTC CF89-12 commenced with an ATG start codon (np 815–817 bp) and terminated with a stop codon (np 2813–2815 bp), and a probable RB site, AGGAG (np 805–809 bp), was also identified to exist immediately upstream of the start codon (AB736316). The putative gene is predicted to encode peptides of 666 amino acid residues with an approximately CMW of 74.0 kDa.

An attempt was then made to analyse the *porA1* gene locus using the PCR experimental procedure with the primer pair MOMPCla_0433F and PorA1gapseqR1 (Figs. 2A and C). As shown in Figure 8, the eight UPTC isolates of the nine (lanes 1–9) generated approximately 4000 bp amplicons and all the 14 UN C. *lari* isolates (lanes 10–23) generated approximately 1800–1900 bp amplicons (Fig. 8). Thus, these eight UPTC isolates may carry the *porA1* gene locus, including the *MCP-dstp* gene, within their genomic DNA.

Nucleotide sequence of the* porA2 *gene locus

During the process of genome sequencing, the *porA2* gene

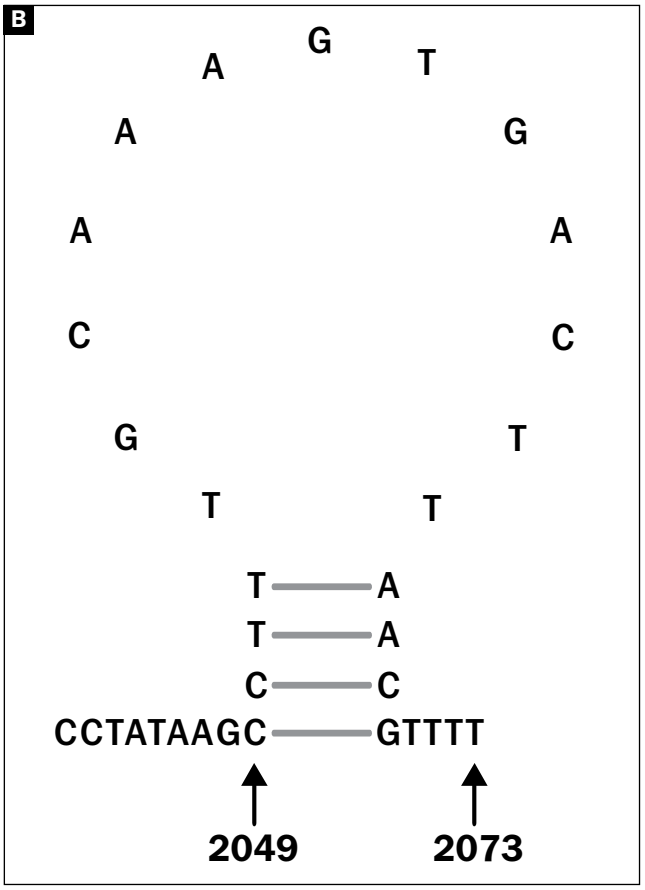


Fig. 6. Nucleotide sequence alignment analysis of the non-coding regions downstream of the *porA1* and *A2* structural genes for the putative transcriptional terminators (A) and their possible secondary structure for *porA2* of the UN *C. lari* JCM2530^T (B). The transcription terminator region for the *porA1* in the UN *C. lari* JCM2530^T is underlined (A).

locus was found within the UPTC CF89-12 and consisted of a partial Cla_1107, an approximate 250 bp NC region, the putative *porA2* structural gene of approximate 1.2 kbp (1206 bp in length; np 474–1679 bp, AB736317), an approximate 290 bp NC region and a partial Cla_1109 structural gene. Interestingly, this *porA2* gene locus (AB736317) is similar to that from the *C. lari* RM2100 (Fig. 2). In addition, as shown in Figure 5, low nucleotide sequence similarity (approximately 57.6%) was seen within the approximate 140 bp NC region upstream of the T-rich region in the *porA1* and *porA2* genes from the UPTC CF89-12 with that in the other six UN *C. lari* isolates.

Table 2. Putative ORFs and their CMWs of *porA1* and *porA2* genes from *C. lari* isolates.

	porA1			*porA2*		
Campylobacter	ORF (bp)	Number of amino acids	CMW (Da)	ORF (bp)	Number of amino acids	CMW (Da)
C. lari JCM2530^T	1233	411	45,667	1233	411	45,667
C. lari 299	1275	425	47,222	1275	425	47,222
C. lari 300	1236	412	45,778	1236	412	45,778
C. lari 448	1275	425	47,222	1275	425	47,222
C. lari 84C-1	1233	411	45,667	1233	411	45,667
C. lari RM2100	1275	425	47,222	1275	425	47,222
UPTC CF89-12	1203	401	44,556	1203	401	44,556

ORF: open reading frame; CMW: calculated molecular weight; Da: Daltons.

Discussion

In the present study, PCR was carried out in order to clone two MOMP structural genes, including putative promoter and terminator structures and their adjacent genetic loci, from the five UN *C. lari* isolates using two primer pairs designed *in silico*. The sequences of approximate 2.1 kbp encoding an approximate 400 bp NC region, possible ORFs of the putative two MOMP genes of approximately 1.2 kbp, approximately 20 bp NC regions and partial and putative CLA_0435 ORF or an NC region of approximately 290 bp and partial and putative Cla_1109 were identified within all five isolates. Each MOMP structural gene (*porA1* and *porA2*) carried start and stop codons and RB sites. The possible ORFs of the two putative MOMP genes were shown to be 1233–1275 bp in length in the five isolate (Table 2).

The nucleotide sequences of the possible ORFs of the putative full-length MOMP structural *porA1* gene from all five UN *C. lari* isolates shared 84.7–99.8% sequence similarity, with *porA2* showing 85.4–99.8% similarity. The nucleotide sequences of the possible ORFs of the putative *porA1* and *porA2* genes from all five UN *C. lari* isolates also shared 84.7–99.9% sequence similarity. In addition, the nucleotide sequences of the MOMP gene (*porA1* and *porA2*) ORFs of the five UN *C. lari* isolates and *C. lari* RM2100 employed as a reference strain shared 55.5–62.7% similarity with three *C. jejuni* NCTC11168 RM1221 and *C. coli* RM2228 strains.

Deduced amino acid sequence alignment analysis was carried out for possible ORFs of the putative full-length MOMP structural genes of the UN *C. lari* isolates, as well as those of *C. lari* RM2100 and three isolates of *C. jejuni* and *C. coli*. Possible ORFs of MOMP genes (*porA1* and *porA2*) from five UN *C. lari* and *C. lari* RM2100 isolates shared 78.4–100% amino acid sequence similarity with each other

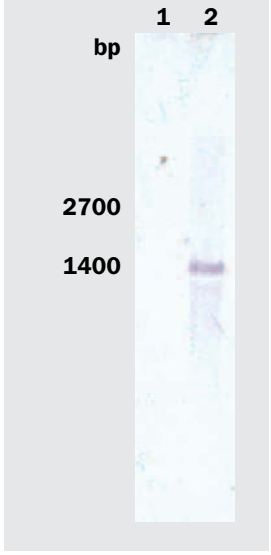


Fig. 7. Northern blot hybridisation analysis of the MOMP structural gene transcript expressed in the UN *C. lari* JCM2530^T. Lane 1: *Escherichia coli* DH5α; Lane 2: UN *C. lari* JCM2530^T.

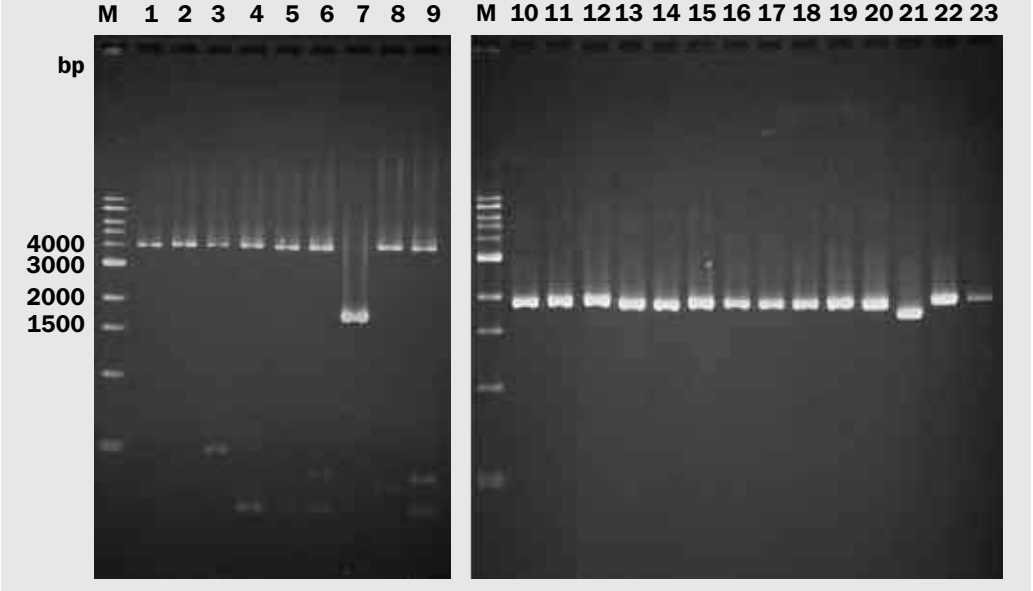


Fig. 8. Agarose gel electrophoresis profiles of the amplicons of the *porA1* locus including the *MCP-dstp* gene with the *C. lari* isolates using the newly designed primer pair, MOMPCla_0433F and PorA1gapseqR1. Lane M: 1 kbp DNA ladder; Lane 1, UPTC A1; Lane 2: 89049; Lane 3: 2; Lane 4: 27; Lane 5: 136; Lane 6: 150; Lane 7: 182; Lane 8: 476; Lane 9: 504; Lane 10: UN *C. lari* JCM2530^T; Lane 11: 28; Lane 12: 170; Lane 13: 175; Lane 14: 176; Lane 15: 264; Lane 16: 274: Lane 17: 295; Lane 18: 298; Lane 19: 299; Lane 20: 300; Lane 21: 382; Lane 22: 448; Lane 23: 84C-1.

Table 3. Nucleotide (upper right) and deduced amino acid (lower left) sequences similarities (%) of full-length *porA1* and *porA2* ORFs from *C. lari* isolates and other thermophilic campylobacters.

	Campylobacter	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	*C. lari* JCM2530^T *porA1*		87.5	94.8	88.8	99.8	87.5	79.5	99.8	88.3	94.9	88.3	99.7	88.5	79.7	62.2	55.6	60.8
2	*C. lari* 299 *porA1*	80.8		87.3	98.2	84.7	99.9	70.1	84.8	98.8	87.4	98.0	84.7	99.0	70.2	59.6	58.8	60.4
3	*C. lari* 300 *porA1*	91.7	82.8		89.4	94.7	90.2	78.5	94.8	89.2	99.9	89.2	94.6	89.3	78.5	62.6	55.7	60.7
4	*C. lari* 448 *porA1*	82.5	97.4	83.7		85.8	98.3	70.5	85.8	99.1	86.8	99.6	85.7	99.2	70.7	60.8	58.9	60.1
5	*C. lari* 84C-1 *porA1*	99.3	78.4	91.7	79.5		87.6	79.6	99.9	88.4	95.0	88.4	99.8	88.6	79.8	62.3	55.6	60.8
6	*C. lari* RM2100 *porA1*	80.8	100	85.4	97.4	81.0		70.2	84.8	98.9	87.5	98.0	84.7	99.1	70.3	59.5	58.8	60.4
7	UPTC CF89-12 *porA1*	74.9	69.4	72.8	70.1	75.2	69.4		81.6	75.0	80.7	74.9	81.4	75.1	99.8	56.6	55.8	63.5
8	*C. lari* JCM2530^T *porA2*	99.5	83.3	92.0	79.8	99.8	78.6	77.1		88.5	95.1	88.5	99.8	88.7	79.8	62.3	55.7	60.8
9	*C. lari* 299 *porA2*	82.0	98.4	83.7	98.6	82.2	98.4	74.1	82.5		86.5	98.8	85.4	99.8	70.9	60.1	57.9	60.0
10	*C. lari* 300 *porA2*	91.7	82.8	100	81.2	92.0	82.8	74.8	92.2	81.2		89.3	94.7	89.4	78.6	62.7	55.8	60.8
11	*C. lari* 448 *porA2*	81.5	97.2	83.5	99.1	81.8	97.2	73.8	82.0	98.1	83.5		85.6	99.0	70.8	60.8	59.0	60.2
12	*C. lari* 84C-1 *porA2*	99.3	78.4	91.7	79.5	99.5	78.4	76.8	99.8	79.5	91.7	79.1		88.5	79.6	62.1	55.5	60.7
13	*C. lari* RM2100 *porA2*	82.2	98.6	84.0	98.8	82.5	98.6	74.3	82.7	99.8	84.0	98.4	82.5		71.0	60.1	58.9	60.3
14	UPTC CF89-12 *porA2*	74.9	69.4	72.8	70.1	75.2	69.4	99.8	75.2	69.9	72.8	69.6	74.9	70.1		56.6	55.9	63.7
15	*C. jejuni* NCTC11168 *porA*	47.4	47.8	48.5	47.8	47.7	47.8	48.4	47.9	47.8	48.5	47.3	47.9	48.0	48.6		80.9	78.9
16	*C. jejuni* RM1221 *porA*	46.7	45.9	47.6	45.2	47.0	45.9	48.4	47.0	45.4	47.6	45.2	47.0	45.6	48.4	73.3		70.4
17	*C. coli* RM2228 *porA*	46.2	45.6	46.8	45.4	46.2	45.6	48.1	46.5	45.4	46.8	45.4	46.5	45.6	48.1	72.4	76.3	

and 45.2–48.5% similarity with those of the ORFs of the three *C. jejuni* and *C. coli* isolates.

Struyve *et al*. showed that the C-terminal Phe is highly conserved in bacterial outer membrane proteins and it plays an important role in the correct assembly of membrane proteins.[29] In the present study, the C-terminal Phe (F) is also conserved in the *porA1* and *porA2* from all six UN *C. lari* and a UPTC isolate, as shown in Figure 9. In addition, the last nine amino acids (VRLQALYKF) at the C-terminus of the MOMPs appeared to be conserved among the strains of *Campylobacter*.[30] Within the seven *C. lari* isolates examined in the present study, identity of the last nine were shown to occur, as indicated in Figure 9. Thus, the last nine amino acids residues could be completely identical among these seven *C. lari* isolates.

Regarding the PorA of *Campylobacter* organisms, the question of whether or not *C. jejuni* produces a cholera toxin-like toxin (CTLT) has been controversial.[31] Albert *et al*. described that CT antibody reacts by immunoblotting with a recombinant PorA, but antibody to the recombinant PorA does not react with CT, and consequently *C. jejuni* does not produce a functional CTLT.[32] Albert *et al*. also described that the PorA proteins from many *Campylobacter* species cross-react with CT.[33]

In addition, the reactivity of PorA with CT antibody would lead to the erroneous conclusion that *C. jejuni* produces a functional CTLT.[31,32] Interestingly, the authors have already shown that *C. lari* UPTC CF89-12 is CT-positive by a Western blot analysis with *Escherichia coli* CT antibody (anti-cholera toxin antibody produced in rabbit, Sigma Louis MO, USA) (data not shown).

The UPTC CF89-12 isolate carried the *porA1* gene locus containing the extra *MCP-dstp* gene (2001 bp in length). Following PCR with the primer pair MOMPCla_0433F and PorA1gapseqR1, the eight UPTC isolates generated approximately 4000-bp amplicons that may carry the extra *MCP-dstp* gene within the *porA1* locus. This MCP (methyl-accepting chemotaxis protein; GenBank ACM64846) is a bacterial transmembrane sensor protein.[34–36] The occurrence of MCPs is widespread throughout the prokaryotes.[35] The *porA1* locus, including the extra *MCP-dstp* gene, may be generated through natural recombination and insertion events within the genomic DNA of the eight UPTC isolates examined.

The *MCP-dstp* gene carried the semi-conserved T-rich region and the promoter consensus sequences at the –10-like region (TAAAAT; np 792–797 bp) and RB site (AGGAG; np 805–809 bp) upstream of the start codon ATG (np 815–817 bp). In addition, as a hypothetically intrinsic transcription terminator structure that contains a G+C-rich

C. jejuni NCTC11168 PorA	412	DHSTVRLQALYKF	424
C. lari JCM2530^T PorA1	399	NDQFI..E.....	411
C. lari JCM2530^T PorA2	399	NDQFI..E.....	411
C. lari 299 PorA1	413	NDQFI..E.....	425
C. lari 299 PorA2	413	NDQFI..E.....	425
C. lari 300 PorA1	400	NDQFI..E.....	412
C. lari 300 PorA2	400	NDQFI..E.....	412
C. lari 448 PorA1	413	NDQFI..E.....	425
C. lari 448 PorA2	413	NDQFI..E.....	425
C. lari 84c-1 PorA1	399	NDQFI..E.....	411
C. lari 84c-1 PorA2	399	NDQFI..E.....	411
C. lari RM2100 PorA1	413	NDQFI..E.....	425
C. lari RM2100 PorA2	413	NDQFI..E.....	425
UPTC CF89-12 PorA1	389	.DQFI..E.....	401
UPTC CF89-12 PorA2	389	.DQFI..E.....	401
		** *****	

Fig. 9. Amino acid sequence alignment analysis of the C terminus of MOMPs among the isolates of the genus *Campylobacter*. Numbers at the left and right refer to the amino acid positions in the possible ORF for each *porA* gene.

Table 4. Nucleotide sequence similarities (%) of an approximate 200 bp of the NC regions upstream of *porA1* and *porA2* ORFs from *C. lari* isolates.

Campylobacter	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1 *C. lari* JCM2530^T *porA1*		96.7	97.7	96.7	99.5	97.2	74.4	99.5	97.2	97.7	96.7	99.5	97.2	74.4
2 *C. lari* 299 *porA1*			98.1	97.2	98.1	99.5	69.5	98.1	99.5	98.1	97.2	98.1	99.5	69.5
3 *C. lari* 300 *porA1*				98.1	99.1	98.6	72.8	99.1	98.6	100.0	98.1	99.1	98.6	72.8
4 *C. lari* 448 *porA1*					97.2	96.7	73.5	97.2	96.7	97.2	100.0	97.2	96.7	73.5
5 *C. lari* 84C-1 *porA1*						97.7	74.9	100.0	97.7	98.1	97.2	100.0	97.7	74.9
6 *C. lari* RM2100 *porA1*							73.2	98.6	100.0	98.6	97.7	98.6	100.0	73.2
7 UPTC CF89-12 *porA1*								73.2	70.9	70.5	71.8	73.2	70.9	100.0
8 *C. lari* JCM2530^T *porA2*									97.7	98.1	97.2	100.0	97.7	74.9
9 *C. lari* 299 *porA2*										98.6	97.7	98.6	100.0	73.2
10 *C. lari* 300 *porA2*											98.1	99.1	98.6	72.8
11 *C. lari* 448 *porA2*												97.2	96.7	73.5
12 *C. lari* 84C-1 *porA2*													97.7	74.9
13 *C. lari* RM2100 p*porA2*														73.2
14 UPTC CF89-12 *porA2*														

NC: non-coding.

region near the base of the stem (np 2824–2845 bp) occurred immediately downstream of the stop codon TAA (np 2813–2815), this *MCP-dstp* gene may be monocistronic.

In conclusion, Southern blot hybridisation analysis suggests two MOMP genes or homologues (*porA1* and *porA2*) in UN *C. lari* organisms. Two MOMP structural genes, with possible ORF of 1233–1275 bp in length, including putative promoter and terminator structures, and their adjacent genetic loci, were molecularly cloned, sequenced and analysed from UN *C. lari* isolates. These genes were confirmed to be transcribed *in vivo* using Northern blot hybridisation analysis with UN *C. lari* JCM2530^T. In addition, eight UPTC isolates may carry the *porA1* gene locus, including the *MCP-dstp* gene, possibly monocistronically, within their genomic DNA. Thus, a genetic heterogeneity occurs within the *porA1* gene locus in some *C. lari* organisms, including UPTC CF89-12. □

This research was partially supported by a project grant awarded 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).

References

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

2 Benjamin J, Leaper S, Owen R J, Skirrow MB. Description of *Campylobacter laridis*, a new species comprising the nalidixic acid resistant thermophilic *Campylobacter* (NARTC) group. *Curr Microbiol* 1993; **8**: 231–8.

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

4 Nachamkin I, Stowell C, Skalina D, Jones AM. *Campylobacter laridis* causing bacteremia in an immunosuppressed patient. *Ann Intern Med* 1984; **101** (1): 55–7.

5 Martinot M, Jaulhac B, Moog R *et al*. *Campylobacter lari* bacteremia. *Clin Microbiol Infect* 2001; **7** (2): 96–7.

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** (3): 1053–5.

7 Bolton FJ, Holt A, Hutchinson DN. Urease-positive thermophilic campylobacters. *Lancet* 1985; **1** (8439): 1217–8.

8 Mégraud F, Chevrier D, Desplaces N, Sedallian A, Guesdon JL. Urease-positive thermophilic *Campylobacter* (*Campylobacter laridis* variant) isolated from an appendix and from human feces. *J Clin Microbiol* 1988; **26** (5): 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** (1): 69–78.

10 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.

11 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** (1): 79–88.

12 Fitzgerald C, Jones K, Anderton S *et al*. Campylobacters in wild birds: identification and molecular characterization. In: Newell DG, Lastovica EE eds. *Campylobacter, Helicobacter and Related Organisms*. Cape Town: Rustica Press, 1988: 80.

13 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** (4): 321–4.

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

15 Dedieu L, Pagès JM, Bolla JM. Environmental regulation of *Campylobacter jejuni* major outer membrane protein porin expression in *Escherichia coli* monitored by using green fluorescent protein. *Appl Environ Microbiol* 2002; **68** (9): 4209–15.

16 Bolla JM, Loret E, Zalewski M, Pagès JM. Conformational analysis of the *Campylobacter jejuni* porin. *J Bacteriol* 1995; **177** (15): 4266–71.

17 Bolla JM, Dè E, Dorez A, Pagès JM. Purification, characterization and sequence analysis of Omp50, a new porin isolated from *Campylobacter jejuni*. *Biochem J* 2000; **352** (Pt 3): 637–43.

18 Dè E, Jullien M, Labesse G, Pages JM, Molle G, Bolla JM. MOMP

(major outer membrane protein) of *Campylobacter jejuni*; a versatile pore-forming protein. *FEBS Lett* 2000; **469** (1): 93–7.
19 Khan I, Adler B, Haridas S, Albert MJ. PorA protein of *Campylobacter jejuni* is not a cytotoxin mediating inflammatory diarrhoea. *Microbs Infect* 2005; **7** (5–6): 853–9.
20 Labesse G, Garnotel E, Bonnel S, Dumas C, Pages JM, Bolla JM. MOMP, a divergent porin from *Campylobacter*: cloning and primary structural characterization. *Biochem Biophys Res Commun* 2001; **280** (1): 380–7.
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** (1): e15.
22 Miller WG, Wang G, Binnewies TT, Parker CT. The complete genome sequence and analysis of the human pathogen *Campylobacter lari*. *Foodborne Pathog Dis* 2008; **5** (4): 371–86.
23 Harrington CS, Thomson-Carter FM, Carter PE. Evidence for recombination in the flagellin locus of *Campylobacter jejuni*: implications for the flagellin gene typing scheme. *J Clin Microbiol* 1997; **35** (9): 2386–92.
24 Sambrook J, Russell DW. *Molecular cloning; a laboratory manual* 3rd edn. New York: Cold Spring Harbor Laboratory Press, 2001.
25 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** (22): 4673–80.
26 Tasaki E, Hirayama J, Tazumi A *et al*. Molecular identification and characterization of clustered regularly interspaced short palindromic repeats (CRISPRs) in a urease-positive thermophilic *Campylobacter* sp. (UPTC). *World J Microbiol Biotechnol* 2012; **28** (2): 713–20.
27 Petersen L, Larsen T S, Ussery DW, On SL, Krogh A. RpoD promoters in *Campylobacter jejuni* exhibit a strong periodic signal instead of a -35 box. *J Mol Biol* 2003; **326** (5): 1361–72.
28 Benjamin L. *Genes VII*. Oxford: Oxford University Press, 2000.
29 Struyvé M, Moons M, Tommassen J. Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein. *J Mol Biol* 1991; **218** (1): 141–8.
30 Zhang Q, Meitzler J C, Huang S, Morishita T. Sequence polymorphism, predicted secondary structures, and surface-exposed conformational epitopes of *Campylobacter* major outer membrane protein. *Infect Immun* 2000; **68** (10): 5679–89.
31 Albert MJ. Cross-reactivity of outer membrane proteins of *Campylobacter* species with cholera toxin. *Indian J Med Res* 2011; **133**: 207–11.
32 Albert MJ, Haridas S, Steer D, Dhaunsi GS, Smuith AI, Adler B. Identification of a *Campylobacter jejuni* protein that cross-reacts with cholera toxin. *Infect Immun* 2007; **75** (6): 3070–3.
33 Albert MJ, Haridas S, Adler B. Major outer membrane proteins from many *Campylobacter* species cross-react with cholera toxin. *Clin Vaccine Immunol* 2008; **15** (6): 859–62.
34 Hagman KE, Porcella SF, Popova TG, Norgard MV. Evidence of a methyl-accepting chemotaxis protein gene (*mcp1*) that encodes a putative sensory transducer in virulent *Treponema pallidum*. *Infect Immun* 1997; **65** (5): 1701–9.
35 Gonzalez I, Richardson PT, Collins MD, Park SF. Identification of a gene encoding a methyl-accepting chemotaxis-like protein from *Campylobacter coli* and its use in a molecular typing scheme for campylobacters. *J Appl Microbiol* 1998; **85** (2): 317–26.
36 Xu Q, Black WP, Nascimi HM, Yang Z. DifA, a methyl-accepting chemoreceptor protein-like sensory protein, uses a novel signaling mechanism to regulate exopolysaccharide production in *Myxococcus xanthus*. *J Bacteriol* 2011; **193** (3): 759–67.