First restriction and genetic mapping of the genomic DNA of urease-positive thermophilic campylobacters (UPTC), and small restriction fragment sequencing

T. ARITOMI*, T. SEKIZUKA*, R. IMAMAKI*, O. MURAYAMA*, B. C. MILLAR†, J. E. MOORE† and M. MATSUDA*

Laboratory of Molecular Biology, School of Environmental Health Sciences, Azabu University, Sagamihara, Japan; and †Department of Bacteriology, Northern Ireland Public Health Laboratory, Belfast City Hospital, Belfast BT9 7AD, Northern Ireland, UK

Accepted: 18 May 2006

ABSTRACT

A restriction and genetic map of urease-positive thermophilic campylobacter (UPTC) CF89-12 genome DNA is constructed using a pulsed-field gel electrophoresis procedure after digestion with *Sal*I and *Sma*I and Southern blot hybridisation. Each of the six gene fragments (*flaA*, *glyA*, *lysS*, *recA*, *sodB* and *ureAB*) selected are mapped in only a fragment on the restriction map. Three DNA fragments for *rrn* operon probes are mapped in multiple regions on the map. When two *Sma*I-digested neighbouring small fragments hybridised with *rrn* probes are cloned and sequenced, a total sequence length of 7487 bp is determined. In the sequence, part of the *pnp* gene (734 bp) bearing a ρ-independent transcriptional termination region, a cluster of five tRNA genes including the putative promoter region, a hypothetical Cj0171-like 507-bp sequence containing an internal termination codon, and a part of the *rrn* operon including the putative promoter region (4700 bp) are identified. The 507 bp sequence carried both putative transcriptional promoter sequences, including a ribosome binding site upstream of the ATG start codon and a characteristic G9 structure, and a possible ρ-independent transcriptional termination region. A hypothetical Cj0170-like 204-bp sequence containing an internal termination codon also occurred, overlapping partly with the Cj0171-like sequence. Based on nucleotide sequence alignment analysis between the UPTC *rrn* operon examined here and the previously reported one, two different 16S-23S ribosomal DNA (rDNA) internal spacer regions are shown to exist.

KEY WORDS: Campylobacter. UPTC.
Chromosome mapping.
Electrophoresis, gel, pulsed-field.
Pseudogenes.
Southern blotting.

Introduction

Campylobacter lari is a thermophilic *Campylobacter* species that was first demonstrated to be resistant to nalidixic acid (nalidixic acid-resistant thermophilic *Campylobacter* sp., NARTC) in 1980.[1,2] Five years later, following the first identification of *C. lari*, urease-positive thermophilic campylobacters (UPTC) were isolated from the natural environment in England.[3,4] Since the first description of UPTC isolates appeared, they have been reported in several European countries[5–10] and in one Asian country.[11,12] The characterisation of UPTC as a variant or a biovar of *C. lari* has been put forward.[4,5]

The authors have already characterised the genomic DNA from UPTC isolates from Japan and Northern Ireland by pulsed-field gel electrophoresis (PFGE) analysis.[9,11] Furthermore, the generation of a restriction and genetic map of UPTC may help to direct cloning and sequencing strategies for this organism.

Therefore, the aim of the current study is to perform restriction and genetic mapping of UPTC in order to elucidate the molecular characteristics of the genomic DNA. In addition, it aims to clone, sequence and characterise some of the small restriction fragments, hybridised with *rrn* operon probes, from the constructed restriction map.

Materials and methods

The UPTC isolate CF89-12 analysed in the present study was isolated from river water in Japan.[11] Culture, intact genomic DNA preparation in an agarose block, and PFGE analysis were carried out, as described previously.[9,11] In order to construct a restriction map, the restriction fragments were fractionated by PFGE after complete, partial and double digestions of the genomic DNA of UPTC CF89-12 with *Sal*I and S*ma*I.

A total of nine housekeeping or representative gene fragments (the *rrn* operon, namely, 16S rDNA, 16S-23S rDNA internal spacer region [ISR], 23S rDNA and, *flaA*, *glyA*, *lysS*, *recA*, *sodB* and *ureAB*]) selected were mapped on the PFGE fragments. Hybridisation probes were prepared according to the procedures described by Sambrook *et al.*,[13]

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

using DIG-High Prime (Roche Diagnostics, Tokyo, Japan) following polymerase chain reaction (PCR) amplification of the genes and fragments, as described elsewhere.[14–19] Southern blot hybridisation for genetic mapping after PFGE was carried out according to the procedure described by Sambrook *et al*.[13]

The resultant recombinant plasmids containing the desired fragments were sequenced using the Texas Red-labelled primer pair of M13. Sequencing of the cloned DNA was performed in an SQ5500EL Hitachi DNA autosequencer (Hitachi Electronics, Tokyo, Japan). Sequencing analysis was carried out using GENETYX-MAC (version 9; Software Development Co., Tokyo, Japan). Nucleotide sequence data obtained in the present study are accessible in the DDBJ/EMBL/GenBank (AB211985).

Results and discussion

Previous PFGE analysis of the genomic DNA of the two Japanese isolates of UPTC CF89-12 and –14 after digestion with *Sal*I and *Sma*I demonstrated that *Sal*I cleaved the genomic DNA from both isolates into six fragments of approximately 960, 460, 210, 135, 70 and 17 kb, and that *Sma*I cleaved the DNA into seven fragments of 870, 460, 210, 135, 110, 85, and 20 kb. In addition, a small fragment of approximately 3.1 kb was observed after digestion with *Sma*I of the genomic DNA from both isolates on conventional agarose gel electrophoresis.[11]

In the present study, a restriction and genetic map of the UPTC CF89-12 genomic DNA was constructed using the PFGE procedure after complete, partial and double digestions with the restriction enzymes *Sal*I and *Sma*I and Southern blot hybridisation (data not shown). Figure 1 shows the restriction map and the resultant genetic map of the restriction fragments. Each of the six genes selected (*flaA*, *glyA*, *lysS*, *recA*, *sodB* and *ureAB*) was mapped to only a fragment of the restriction map (the *Sal*I-5/*Sma*I-7 region for *flaA*, *recA*, *sodB* and *ureAB* and the *Sal*I-2/*Sma*I-3 region for *glyA* and *lysS*).

In contrast, three *rrn* fragments (16S rDNA, 16S-23S rDNA ISR and 23S rDNA) were mapped in multi-fragments on the map (i.e., *Sal*I-1 *Sal*I-3/*Sma*I-5 *Sma*I-6 region, the *Sal*I-4/*Sma*I-4 region and the *Sal*I-2 *Sal*I-4/*Sma*I-1 *Sma*I-2 *Sma*I-3 region). Thus, the present study results may provide evidence that at least the three *rrn* operons exist in the genome of this UPTC isolate.

An attempt was made to clone and sequence the two *Sma*I-digested small fragments, *Sma*I-1 (approximately 3.1 kb) and a part of *Sma*I-2, both hybridised with *rrn* operon probes (16S rDNA, 16S-23S rDNA ISR and 23S rDNA) using a pUC19 vector, in an effort to elucidate the molecular characteristics of the genetic information in the two small fragments.

First, the *Sma*I-1 fragment was digested with restriction enzymes (*Bam*HI, *Xba*I, *Sph*I and *Sal*I) and the restriction map was constructed (Fig. 2A). The resultant fragments were transformed in *Escherichia coli* DH5α cells and then sequenced. Restriction maps were constructed for the *Sma*I-2 fragment with *Pst*I (Fig. 2B) and for the *Sma*I-2 (*Pst*I-1) of the 4.4 kb fragment with *Hind*III and *Kpn*I (Fig. 2B). The resultant fragments were transformed in *E. coli* DH5 α cells and sequenced.

Consequently, in the present study a total region of approximately 7.5 kb in length was cloned and sequenced, comprising the two fragments neighbouring each other of approximately 3.1 kb for the *Sma*I-1 fragment and 4.4 kb for the *Sma*I-2 (*Pst*I-1) fragment.

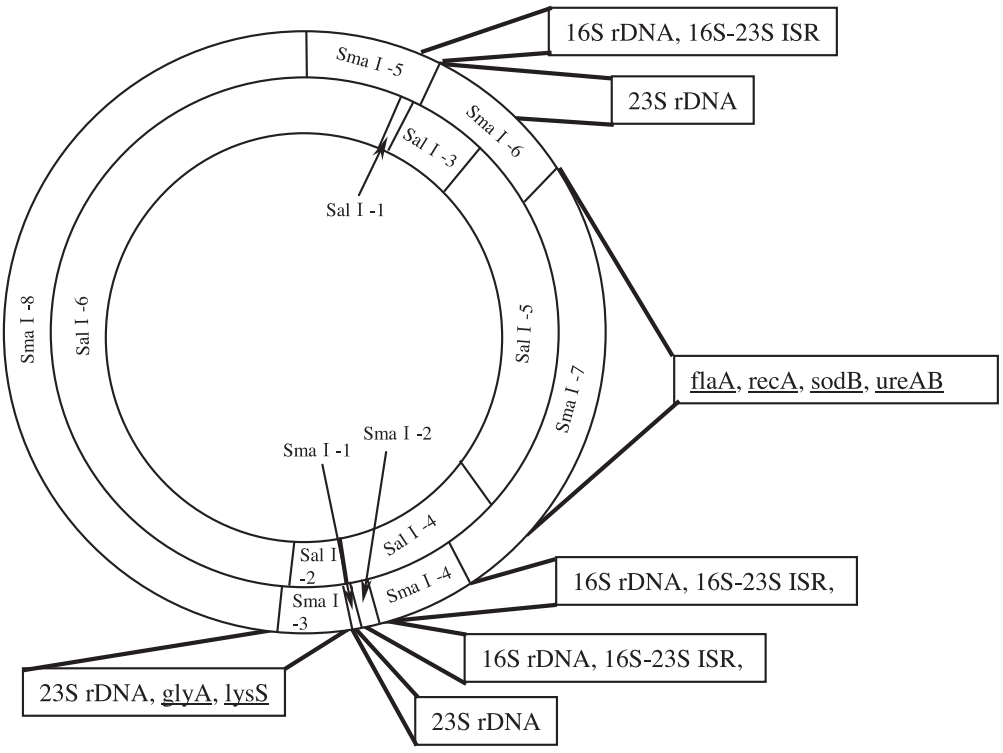
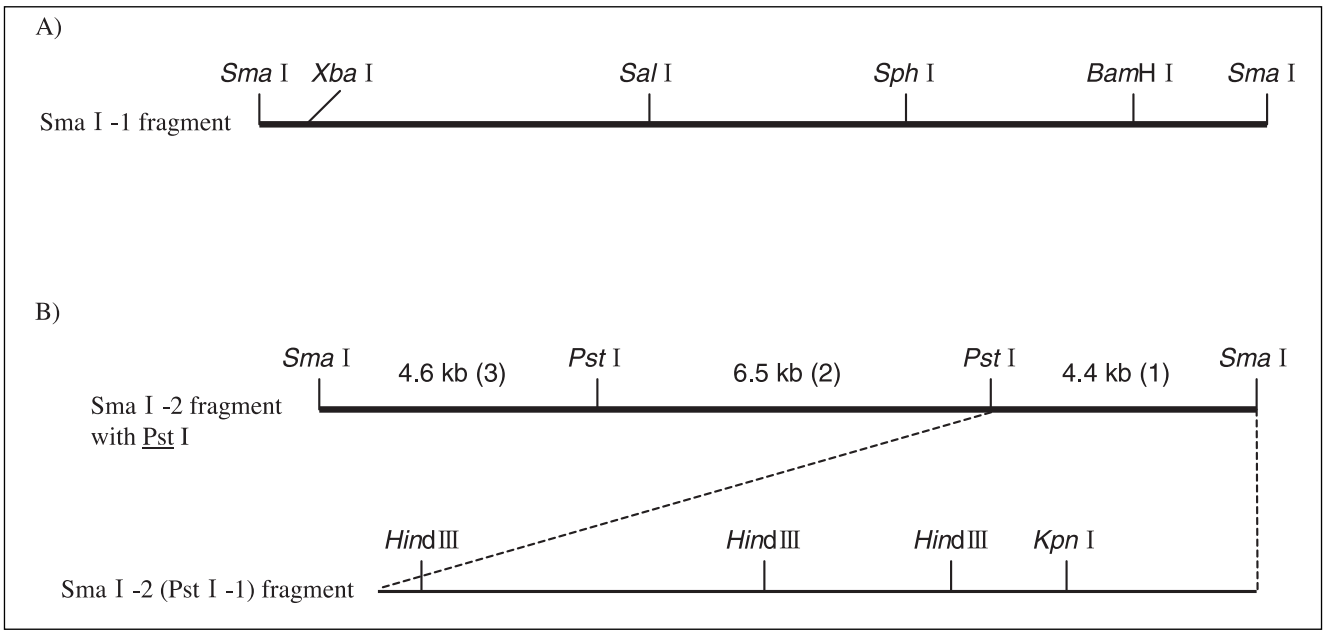


Fig. 1. A restriction and genetic map of UPTC CF89-12 genomic DNA made using the PFGE procedure after digestions with *Sal*I and *Sma*I and Southern blot hybridisation. After construction of a restriction map of CF89-12, a total of nine housekeeping or representative gene fragments (the *rrn* operon [16S rDNA, 16S-23S rDNA ISR, 23S rDNA and, *flaA*, *glyA*, *lysS*, *recA*, *sodB* and *ureAB*]) were mapped on the PFGE fragments.

Fig. 2. Construction of the restriction maps for the *Sma*I-1 fragment with the restriction enzymes, *Bam*H1, *Xba*I, *Sph*I and *Sal*I (A) and for the *Sma*I-2 (*Pst*I-1) fragment (approximately 4.4 kb fragment) with *Hind* III and *Kpn* I (B).



A schematic representation of the genetic organisation that was cloned, sequenced and clarified in a total region of approximately 7.5 kb of two *Sma*I-digested neighbouring small fragments is provided in Figure 3.

From the SmaI-1 fragment (approximately 3.1 kb), a total of 3330 bp (nucleotide position [np] 4158–7487) was cloned, sequenced and identified to contain the 3′ end region of 16S rDNA (150 bp), 16S-23S rDNA ISR (720 bp) and 23S rDNA (2460 bp), based on nucleotide sequence alignment analysis. In addition, from the SmaI-2 (P*st*I-1) fragment (approximately 4.4 kb), a total of 4157 bp was cloned, sequenced and identified to contain the 3′ end region of the polyribonucleotide nucleotidyltransferase (*pnp*) gene (734 bp; np 1–734). It exhibited a high nucleotide sequence similarity to the *pnp* gene from *C. jejuni* (DDBJ/EMBL/GenBank accession No. NC_002163; AAFK01000006)[20–22] and two putative promoter structures consisting of consensus sequences at the –35 (TTGACA; np 779-784) and –10 region (CATAAT; np 802–807).

Moreover, the approximate 4.4 kb fragment contained a five tRNA gene cluster, namely the $tRNA^{Lys}$ (np 822–897), $tRNA^{Val}$ (np 913–988), $tRNA^{Asp}$ (np 1008–1084), $tRNA^{Lys}$ (np 1087–1162) and $tRNA^{Glu}$ genes (np 1169–244). The present tRNA gene cluster contained non-coding regions of 15 bps between the first $tRNA^{Lys}$ and the second $tRNA^{Val}$, 19 $tRNA^{Val}$ $tRNA^{Asp}$, two $tRNA^{Asp}$ $tRNA^{Lys}$ and six $tRNA^{Lys}$ $tRNA^{Glu}$, respectively. Nucleotide sequences of the two $tRNA^{Lys}$ genes in the cluster were identical.

The *pnp* gene also carried a possible ρ-independent transcriptional termination region (np 736–768) downstream of the stop codon UAA (np 732–734). In addition, a 507 bp sequence (np 1553–2059) with high nucleotide sequence similarity to the *C. jejuni* hypothetical protein Cj0171 of unknown function (NC_002163) and a 16S rDNA sequence (1513 bp; np 2794–4306), including the putative promoter sequences, were identified.

Two putative promoter sequences consisting of consensus sequences at the –35 region (TTAAAC; np 1483–1488) and –10 region (TTGATT; np 1508–1513) and ribosome binding (RB) site (TGGAGT; np 1527–1532) exist upstream of the ATG start codon (np 1553–1555) for the sequence of the hypothetical Cj0171 with a characteristic G9 structure (5′-GGGGGGGGG-3′; np 1560–1568). The 507 bp sequence was also followed by a possible ρ-independent transcriptional termination region (np 2068–2094) downstream of the stop codon UAA (np 2057–2059).

Surprisingly, however, the 507 bp sequence contained an internal termination codon UAA (np 1556–1558) following

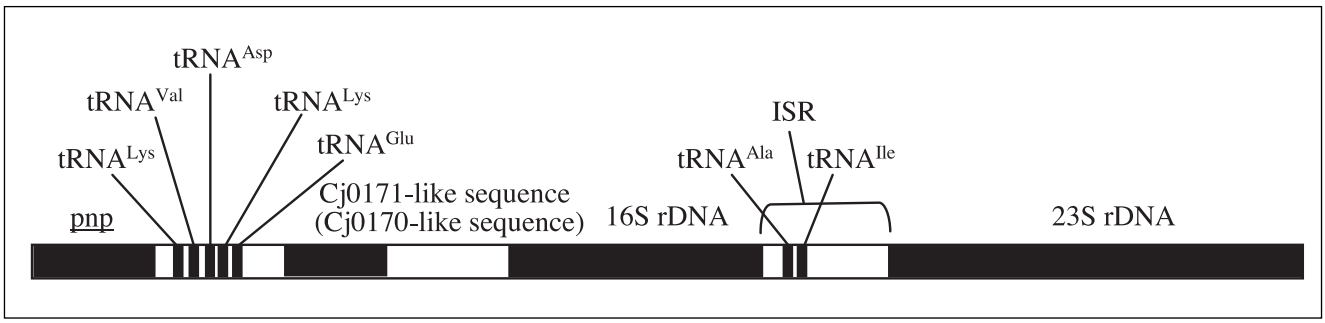


Fig. 3. Schematic representation of the genetic organisation cloned, sequenced and clarified in a total length of approximately 7.5 kb of two *Sma*I-digested neighbouring small fragments.

the start codon (np 1553–1555), which suggests that its incomplete or pseudogene sequence would be non-functional. Although a 204 bp sequence (np 1372–1575) with high nucleotide sequence similarity to the *C. jejuni* hypothetical protein Cj0170 of unknown function (NC_002163), as well as two putative promoter sequences, TATAAT (np 1314–1319) and TATAAT (np 1335–1340) and RB site (AGGAGA; np 1362–1367) was identified, the sequence also contained an internal termination codon UAA (np 1393–1395). An overlap was detected from np 1553 to 1575 between these two sequences.

Thirteen complete repeat units (np 2243–2385) of 11 bp (5′-TTCTTTCTCTA-3′) were characteristically identified between the hypothetical Cj0171-like 507 bp sequence and the two putative promoter sequences (TTAATA at the –35 region [np 2626–2631]; TATAGT at the –10 region [np 2647-2652]) upstream of the 16S rDNA sequences. At present, however, the function of the repeat units is unknown. Moreover, a putative box A-like antitermination element sequence (CTTTTTCTTTGAAGAAA; np 2758–2774) was also identified.

In relation to the nucleotide sequences of the *rrn* operon from UPTC CF89-12, this group has already sequenced and reported a 2423 bp sequence consisting of the 16S rDNA (approximately 1500 bp), 16S-23S rDNA ISR (720 bp) and 23S rDNA (190 bp) by a TA cloning and sequencing procedure (AB066098).[16] When the present nucleotide sequence of the full length 16S rDNA (approximately 1510 bp), 16S-23S rDNA ISR (720 bp) and 23S rDNA (2,450) was compared with those of the previous sequence data described above, it was demonstrated that the corresponding sequences of 16S rDNA and 23S rDNA regions were almost identical (AB06098 and AB211985). However, nucleotide sequence differences were found at a total of 15 loci in the 720 bp sequences of the 16S-23S rDNA ISR between the present and previous sequence data (Fig. 4), indicating two ISRs from the two distinctly different *rrn* operons.

In the present study, an entire five tRNA gene cluster operon was isolated. However, an expression study is required in order to clarify whether or not the operon is functional in UPTC cells.

In relation to the 507 bp and alternative 204 bp sequences with high sequence similarity to Cj0171 and Cj0170, respectively, this is the second demonstration of the incomplete or pseudegene sequences that contain an internal termination codon in a Japanese UPTC isolate CF89-12, following the first *flaA*-like sequences described previously.[23]

In light of these findings, more incomplete or pseudogene sequences may be identified in the UPTC genome, as described already in other campylobacters.[20, 22]

JEM was supported by the Research and Development Office, Department of Health, Northern Ireland (Infectious Disease-Recognised Research Group [RRG] 9.9).

References

1 Skirrow MB, Benjamin J. Differentiation of enteropathogenic *Campylobacter. J Clin Pathol* 1980; **33:** 1122.
2 Benjamin J, Leaper S, Owen RJ, Skirrow MB. Description of *Campylobacter laridis*, a new species comprising the nalidixic acid-resistant thermophilic *Campylobacter* (NARTC) group. *Curr Microbiol* 1983; **8**: 231–8.
3 Bolton FJ, Holt AV, Hutchinson DN. Urease-positive thermophilic campylobacters. *Lancet* 1985; **i**: 1217–8.
4 Owen RJ, Costas M, Sloss L, Bolton J. Numerical analysis of

```
4299  CACCTCCTTTCTAGAGTACAAACTGATAAGTCTCACAACTATCAGTTCATATATAGACTCAATCATCCTTGTTTAGATTTCAAAGATTGATGAAAAGCTA
      ****************************************************************************************************
1499  CACCTCCTTTCTAGAGTACAAACTGATAAGTCTCACAACTATCAGTTCATATATAGACTCAATCATCCTTGTTTAGATTTCAAAGATTGATGAAAAGCTA

4399  ATTTTGGGGAATTAGCTCAGCTGGGAGAGCGCCTGCTTTGCACGCAGGAGGTCAGCGGTTCGATCCCGCTATTCTCCACCATTTATTAAGGGCCTATAGC
      ****************************************** ************************************* ***************
1599  ATTTTGGGGAATTAGCTCAGCTGGGAGAGCGCCTGCTTTGCACACAGGAGGTCAGCGGTTCGATCCCGCTATTCTCCACCATTT-TTAAGGGCCTATAGC

4499  TCAGCTGGTTAGAGTGCACCCCTGATAAGGGTGAGGTCACAAGTTCAAGTCTTGTTAGGCCCACCAT-AAAAAAGATTTGTTTATCAATCATATAATAAA
      ****************************************************************** ********************************
1698  TCAGCTGGTTAGAGTGCACCCCTGATAAGGGTGAGGTCACAAGTTCAAGTCTTGTTAGGCCCACCATAAAAAAAGATTTGTTTATCAATCATATAATAAA

4598  GTTTTAAAACTTAAAGCAAGTATATAGCTAATAATAAAGCAGTAAATAAAGTAAAATACTTTATTGCTTATATTTTTAAATTATTTAAACTCTTTATATA
      *************************** ******** **** ***  * ****************  *********************************
1798  GTTTTAAAACTTAAAGCAAGTATATAGTTAATAATATAGCAATAAGCAGAGTAAAATACTTTATTATTTATATTTTTAAATTATTTAAACTCTTTATATA

4698  TACTTCCTTTAGGTTTTAAGACCTAAGCTAAATAAATAATAGAATATTTTATATATTAGATTAGTTATTTAATGTTATTTAAATTATCATTGTTAAGAGT
      ************************************************************************************ ******************
1898  TACTTCCTTTAGGTTTTAAGACCTAAGCTAAATAAATAATAGAATATTTTATATATTAGATTAGTTATTTAATGTTATTTGAATTATCATTGTTAAGAGT

4798  CACAAGCAAGTTTTAATAAAAACAATTTTACAGGACTTGTTAAAGGATAAAACCTAACTATCTTTTCTTTAGCTTTTAGTTAAAGACAAGTTTTAAACTC
      ****************************************************************************************************
1998  CACAAGCAAGTTTTAATAAAAACAATTTTACAGGACTTGTTAAAGGATAAAACCTAACTATCTTTTCTTTAGCTTTTAGTTAAAGACAAGTTTTAAACTC

4898  ACAACTTAGTTAGACTAATATTTTTAGTTTTATTAAGTGTTTAAATGCTTTCCGTCTTAAGATAGTAAAGTCTTATCATAAAAACTTTAACAAGGAAGTG
      ****************** * ********************** ****************************************************
2098  ACAACTTAGTTAGACTAATTTGTTTAGTTTTATTAAGTGTTTAAAAGCTTTCCGTCTTAAGATAGTAAAGTCTTATCATAAAAACTTTAACAAGGAAGTG
```

Fig. 4. Nucleotide sequence alignment between the present (upper sequence) and the previous (lower) 16S-23S rDNA ISRs. Approximately 700 bp sequences downstream of the CCTCCT sequence (np 4301–4306 for the present ISR and np 1501–1506 for the previous ISR) are shown. Numbers at the left refer to the np of the sequences. Positions identical in both sequences are marked by asterisks.

electrophoretic protein patterns of *Campylobacter laridis* and allied thermophilic campylobacters from the natural environment. *J Appl Bacteriol* 1988; **65**: 69–78.

5 Megraud F, Chevrier D, Desplaces N, Sedallian A, Guesdon J L. Urease-positive thermophilic *Campylobacter* (*Campylobacter laridis* variant) isolated from an appendix and from human faeces. *J Clin Microbiol* 1988; **26**: 1050–1.

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

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

8 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.

9 Kaneko A, Matsuda M, Miyajima M, Moore JE, Murphy PG. Urease-positive thermophilic strains of *Campylobacter* isolated from seagulls (*Larus* spp.). *Lett Appl Microbiol* 1999; **29**: 7–9.

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

11 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.

12 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; **250**: 321–4.

13 Sambrook J, Fritsch E F, Maniatis T. *Molecular cloning, a laboratory manual* (2nd edn). New York: Cold Spring Harbor Laboratory Press, 1989.

14 Newnham E, Chang N, Taylor DE. Expanded genomic map of *Campylobacter jejuni* UA580 and localization of 23S ribosomal rRNA genes by I-*Ceu*I restriction endonuclease digestion. *FEMS Microbiol Lett* 1996; **142**: 223–9.

15 Al Rashid ST, Dakuna I, Louie H *et al*. Identification of *Campylobacter jejuni*, *C. coli*, *C. lari*, *C. upsaliensis*, *Arcobacter butzleri*, and *A. butzleri*-like species based on the *glyA* gene. *J Clin Microbiol* 2000; **38**: 1488–94.

16 Miyajima M, Matsuda M, Haga S, Kagawa S, Millar BC, Moore JE. Cloning and sequencing of 16S rDNA and 16S-23S rDNA internal spacer region (ISR) from urease-positive thermophilic *Campylobacter* (UPTC). *Lett Appl Microbiol* 2002; **34**: 287–9.

17 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.

18 Honda M, Matsushita S, Murayama O, Millar BC, Moore JE, Matsuda M. Cloning and sequence analysis of the *recA* gene in urease-positive thermophilic campylobacters (UPTC). *Br J Biomed Sci* 2002; **59**: 166–9.

19 Matsuda M, Usui K, Iida H *et al*. Cloning and characterization of a urease gene operon in urease-positive thermophilic *Campylobacter* (UPTC). *Int J Hygiene Environ Health* 2006. In press.

20 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.

21 Regonesi ME, Briani F, Ghetta A *et al*. A mutation in polynucleotide phosphorylase from *Escherichia coli* impairing RNA binding and degradosome stability. *Nucleic Acids Res* 2004; **32**: 1006–17.

22 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.

23 Sekizuka T, Gondo T, Murayama O, Moore JE, Millar BC, Matsuda M. *flaA*-like sequences containing internal termination codons (TAG) in urease-positive thermophilic *Campylobacter* isolated in Japan. *Lett Appl Microbiol* 2002; **35**: 185–9.