JOURNAL OF BACTERIOLOGY, May 1993, p. 3051–3057
0021-9193/93/103051-07$02.00/0

Vol. 175, No. 10

Distribution and Polymorphism of the Flagellin Genes from Isolates of *Campylobacter coli* and *Campylobacter jejuni*

RICHARD A. ALM,[1] PATRICIA GUERRY,[2] AND TREVOR J. TRUST[1]*

Department of Biochemistry and Microbiology, University of Victoria, Victoria, British Columbia, V8W 3P6, Canada,[1] and Enteric Diseases Program, Naval Medical Research Institute, Bethesda, Maryland 20814[2]

Received 2 December 1992/Accepted 14 March 1993

The complex flagellar filaments of the LIO8 serogroup member *Campylobacter coli* VC167 are composed of two highly related subunit proteins encoded by the *flaA* and *flaB* genes which share 92% identity. Using oligonucleotide primers based on the known DNA sequence of both the *flaA* and *flaB* genes from *C. coli* VC167 in the polymerase chain reaction, we have shown conservation of both *fla* genes among isolates within the LIO8 heat-labile serogroup by digestion of the amplified product with *Pst*I and *Eco*RI restriction endonucleases. Amplification and subsequent restriction analysis of the *flaA* flagellin gene from *Campylobacter* isolates belonging to 13 different LIO serogroups further identified 10 unique polymorphic groups. Within most of the serogroups examined, isolates appeared to contain *flaA* genes with conserved primary structures. Only in serogroups LIO11 and LIO29 did independent isolates possess flagellin genes with different primary structures. Furthermore, by employing primers specific for the *flaB* gene of *C. coli* VC167, all serogroups examined contained a second *fla* gene corresponding to *flaB*. In all serogroups except the LIO5 and LIO6 isolates which were identical to each other, the polymorphic pattern of this *flaB* gene was identical to that of the corresponding *flaA* gene. These data indicate that the presence of a second highly homologous flagellin gene is widespread throughout *Campylobacter* isolates and that in most instances, the primary structure of the two *fla* genes is conserved within isolates belonging to the same heat-labile LIO serogroup. This may represent the presence of clonal evolutionary groups in *Campylobacter* spp.

The best-described flagella at both the structural and molecular levels are those of the enteric *Escherichia* and *Salmonella* spp. The unsheathed flagella of these organisms are formed by the polymerization of a single species of flagellin. Another group of bacteria produce unsheathed flagellar filaments composed of two or more species of flagellin. These "complex" filaments have been described for bacterial species such as *Campylobacter coli* (14), *Campylobacter jejuni* (27), *Helicobacter pylori* (25), *Caulobacter crescentus* (10), *Bdellovibrio bacteriovorus* (42), and *Rhizobium meliloti* (35). The periplasmic flagella of *Spirochaeta aurantia* (32) and *Treponema pallidum* (31) are also composed of multiple species of flagellin subunits, as are the flagellar filaments of archaea such as *Halobacterium halobium*, *Methanococcus* spp., and *Methanospirillum* spp. (13, 22, 23, 41). For several of these organisms, the flagellin genes coding for the two or more flagellin species have been sequenced. However, as a group, the complex flagella remain poorly described. For example, in most cases the information available is restricted to a single strain of the species, and little is known regarding the antigenic and structural diversity of flagellins which form complex flagella.

Regardless of whether the flagellar filament is simple or complex, the structures of most bacterial flagellin subunits are strikingly similar. For a large range of organisms, the amino- and carboxy-terminal domains of flagellin molecules which have roles in flagellin export, polymerization of the flagellar filament, and filament stability (11, 46, 47) display significant amino acid sequence homology (21, 45, 49). However, the central region of the flagellin molecule often displays large differences in both size and amino acid content and appears to be responsible for the extensive antigenic diversity seen among flagellins from members of the family *Enterobacteriaceae* (18–20, 26). Preliminary immunological evidence suggests that such a model may also be valid for *Campylobacter* spp. (28). Moreover, genetic studies by Thornton et al. (43) demonstrated that serotype-specific DNA probes can be constructed by using the central region of the *Campylobacter* flagellin gene.

The flagellar filament of *C. coli* VC167 can undergo antigenic variation between two antigenic forms, T1 and T2 (15, 17), each of which contains two flagellin subunits, FlaA and FlaB. These flagellins are coded for by two tandemly orientated, highly homologous genes which show 92.0% identity in VC167-T1 and 91.6% identity in VC167-T2 (14, 16). *C. jejuni* 81-116 also possesses two tandem flagellin genes which display 92.8% identity with each other (27), and the total identity between the flagellin genes of strains VC167 and 81-116 is 82%. The *flaA* genes of two other *C. jejuni* strains, IN1 and TGH9011, have also been sequenced (12, 24) and show 74.9 and 73.6% identity to the *flaA* gene of VC167, respectively. This suggests that there may be considerable genetic diversity among the *fla* genes in *Campylobacter* spp. Although a number of studies have suggested the presence of a second *fla* gene in other strains of *Campylobacter* (12, 43), it is not known how widespread *flaB* is throughout *Campylobacter* spp., as tandem gene duplications are often rapidly eliminated by recombination (7, 44), or indeed how related *flaB* is to *flaA*.

One of the major serotyping schemes of *Campylobacter* spp. is the heat-labile scheme of Lior (26a), which recognizes more than 100 serogroups. It had previously been suggested that flagella were the serodeterminant of this scheme, but recent genetic data have indicated that in most serogroups flagella are not the Lior serodeterminant (3). Nonetheless, surface-exposed serospecific determinants on flagella have been demonstrated and flagellin genes are highly conserved

* Corresponding author.

TABLE 1. Strains used in this study

Serogroup	Strain(s)	Polymorphism group	
		flaA	*flaB*
LIO8	*C. coli* VC20, VC143, VC144, VC167, VC189; *C. jejuni* VC156, VC159	1	1
LIO8/(29)	*C. jejuni* VC152, VC157	1	1
LIO6	*C. jejuni* VC84, VC209, VC223, 81-116	2	3
LIO4	*C. jejuni* VC83, VC207, VC226, VC331, VC237	4	4
LIO15	*C. jejuni* VC94, VC185, VC220, VC227	5	5
LIO11	*C. jejuni* VC74	5	5
LIO11	*C. jejuni* VC91	6	6
LIO11	*C. jejuni* VC228, VC232	7	7
LIO11	*C. jejuni* VC230	4	4
LIO19	*C. jejuni* VC104, VC221, VC233	8	8
LIO12	*C. coli* VC92, VC229	9	9
LIO5	*C. jejuni* VC88, 81-176	2	3
LIO20	*C. coli* VC97, VC225	8	8
LIO29	*C. coli* VC168	5	5
LIO29	*C. coli* VC236	10	10
LIO1	*C. jejuni* VC87	7	7
LIO7	*C. jejuni* VC95	11	11
LIO55	*C. coli* VC235	5	5

within Lior serogroups (4, 17). In this study, we undertook to determine the conservation of the primary flagellin gene structure within various LIO serogroups and the polymorphism of the flagellin genes among serogroups. Further, we could determine the presence or absence of *flaB* in a range of *Campylobacter* strains belonging to a variety of LIO serogroups, and, by restriction endonuclease digestion of polymerase chain reaction (PCR)-amplified products that represent ~85% of the flagellin genes, we could investigate the similarity of the primary structure of the *flaB* gene, if present, to that of *flaA*.

MATERIALS AND METHODS

Bacterial strains and culture conditions. The *C. coli* and *C. jejuni* strains listed in Table 1 were obtained from H. Lior, National Enteric Reference Centre, Ottawa, Canada. *Campylobacter* spp. were grown on Mueller-Hinton agar (Difco) at 37°C in an atmosphere containing 5% oxygen and 10% CO_2.

Preparation of chromosomal DNA. A plate of fresh overnight *Campylobacter* growth was harvested in 2.5 ml of ice-cold 25% sucrose–50 mM Tris-HCl, pH 8.0. One milliliter of lysozyme (10 mg/ml in 0.25 M EDTA, pH 8.0) solution was added, and incubation was continued on ice for 20 min. This was followed by the addition of 0.75 ml of TE (10 mM Tris-HCl, 1 mM EDTA, pH 8.0), 0.25 ml of lysis solution (5% *N*-lauroylsarcosine; 50 mM Tris-HCl, pH 8.0; 62.5 mM EDTA), and 10 mg of pronase (Boehringer-Mannheim Canada Ltd., Laval, Québec, Canada), and incubation was continued at 58°C for 90 min. The viscous DNA solution was then extracted twice with Tris-saturated phenol and once with diethyl ether before being placed in dialysis bags and dialyzed against 2 liters of TE at 4°C for 16 to 24 h with one change.

PCR amplification. Oligonucleotides used for amplification were constructed on an Applied Biosystems 392 DNA synthesizer and were based on the known sequences of the flagellin genes of *C. coli* VC167. The *flaA*-specific 18-bp primer pg50 (5′-ATGGGATTTCGTATTAAC-3′) (30) used for PCR amplification is located on the plus strand beginning at the N terminus (Fig. 1A). The *flaB*-specific 32-bp primer RAA9 (5′-AAGGATTTAAAATGGGTTTTAGAATAAAC ACC-3′) is also located on the plus strand and begins 11 bp upstream of the translational start site of *flaB* (Fig. 1A). The 26-bp primer RAA19 [5′-GCACC(CT)TTAAG(AT)GT(AG) GTTACACCTGC-3′], which binds to both flagellin genes 274 bp from the 3′ end, was used as the reverse primer in both cases (Fig. 1A), and PCR amplification generated the fragments depicted in Fig. 1B. Other primers used were RAA10 (5′-TCTTGCTTTAATCTTTTCGATGCA-3′), which binds between 10 and 33 bp upstream of RAA9 on the same strand, and RAA5 (5′-TTGCACAGCGTTACGTTGGCT-3′), which binds to the *flaB* coding region 25 bp from the translational stop codon on the minus, or noncoding, strand. AmpliTaq polymerase (Perkin-Elmer Cetus, Rexdale, Ontario, Canada) (0.4 U/10 μl of reaction mixture) was used in a hot-air thermocycler (Idaho Technology, Idaho Falls, Idaho) in the presence of 1× standard PCR buffer, 2.5 mM $MgCl_2$ (GeneAmp; Perkin-Elmer Cetus), a final concentration of 250 μM for each primer, and 50 ng of genomic template. Thirty to 40 cycles of amplification were performed in thin-walled capillary tubes, each cycle consisting of 1 s at 94°C, 1 s at 37°C

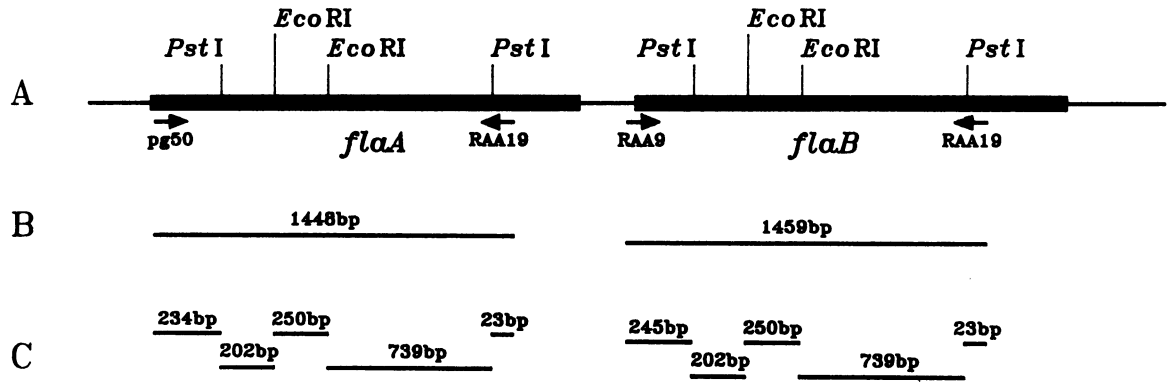

FIG. 1. Flagellin genes from *C. coli* VC167. (A) Schematic representation indicating the localization of the priming sites for the oligonucleotides used for PCR amplification and the restriction endonuclease sites used. (B) PCR products generated with primer pairs pg50-RAA19 and RAA9-RAA19 that represent the majority of the *flaA* and *flaB* genes, respectively. (C) Restriction fragments generated upon digestion of the amplified product with *Pst*I and *Eco*RI.

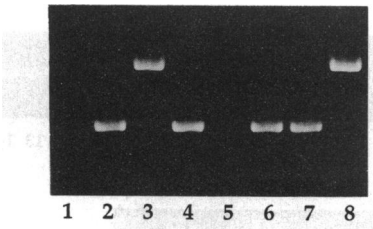

FIG. 2. PCR amplification with primer pairs pg50-RAA19 (lanes 1 to 4) and RAA9-RAA19 (lanes 5 to 8) on strains VC167-T2 (lanes 2 and 6), the *flaA flaB*$^+$ mutant KX15 (14) (lanes 3 and 7), and the *flaA*$^+$ *flaB* mutant KX5 (14) (lanes 3 and 7). No DNA controls are seen in lanes 1 and 5. Fragment sizes are 1,448 bp (lanes 2 and 4), 2,848 bp (lane 3), 1,459 bp (lanes 6 and 7), and 2,859 bp (lane 8).

(*flaA* primers) or 55°C (*flaB* primers), and a 60-s extension at 74°C.

Restriction analysis of PCR products. Following PCR amplification, the 10-μl reaction mixtures were either analyzed directly on 1% agarose gels and visualized by staining with ethidium bromide or subjected to restriction enzyme digestion. PCR products (10 μl) were digested in a total volume of 12 μl with 3 U of both *Pst*I and *Eco*RI for 90 min at 37°C. After digestion, the fragments were separated on 12% polyacrylamide gels in a minislab gel apparatus (Hoefer Scientific Instruments, San Francisco, Calif.) in 1× TBE buffer by the method of Sambrook et al. (36) and visualized by staining with ethidium bromide.

RESULTS

Specificity of PCR primers in LIO8 strains. Because the *flaA* and *flaB* genes of *C. coli* VC167 are so highly homologous, the ability of each primer pair to specifically amplify either flagellin gene had first to be determined. To accomplish this, we employed the previously described flagellin mutants KX5 (*flaA*$^+$ *flaB*) and KX15 (*flaA flaB*$^+$) (14). The *flaA* primer pair (pg50-RAA19) amplified a 1,448-bp product from wild-type VC167-T2 (Fig. 2, lane 2) and the *flaA*$^+$ *flaB* mutant KX5 (Fig. 2, lane 4). A larger product was amplified from the KX15 mutant because of its *flaA* gene being inactivated by a Km^r cassette (Fig. 2, lane 3). Similarly, the *flaB* primer pair (RAA9-RAA19) amplified a 1,459-bp product from VC167-T2 (Fig. 2, lane 6) and KX15 (Fig. 2, lane 7). The expected larger fragment due to the Km^r cassette was amplified from mutant KX5 (Fig. 2, lane 8). These results indicated that in this strain, these primer pairs could specifically amplify the two flagellin genes.

Presence and conservation of *fla* genes among LIO8 isolates. Genomic DNA from the seven LIO8 and two LIO8/(29) strains listed in Table 1 were subjected to PCR amplification with specific *flaA* and *flaB* primer pairs. Certain *Campylobacter* isolates agglutinate with both the LIO8 and LIO29 typing sera and are classed as serogroup LIO8/(29). The *flaA* and *flaB* genes amplified from the LIO8 and LIO8/(29) strains possessed the same size as that seen with VC167 (Fig. 3A and B). To determine whether the primary structure of the flagellin genes was conserved within the LIO8 serogroup, the PCR products were digested with *Pst*I and *Eco*RI, which cut within the central region of the VC167 flagellin genes (Fig. 1A). The sizes of the fragments based on the known sequences of the flagellin genes from VC167 (14, 16) are indicated schematically in Fig. 1C. Indeed, when the PCR products that represented the LIO8 and LIO8/(29) *flaA* and *flaB* genes were digested, they displayed perfect conservation with the VC167-T2 control (Fig. 4A and B).

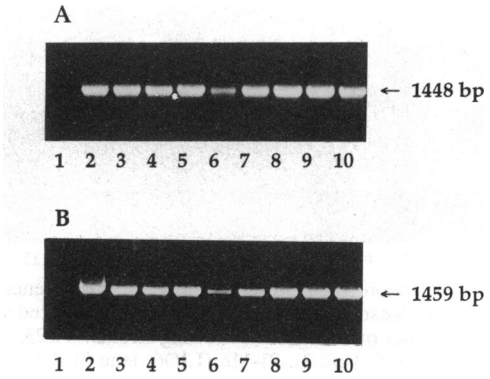

FIG. 3. PCR amplification with primer pairs pg50-RAA19 (A) and RAA9-RAA19 (B) on LIO8 and LIO8/(29) strains listed in Table 1. Lane 1 is the no-DNA control. Strains VC20 (lane 2), VC143 (lane 3), VC144 (lane 4), VC152 (lane 5), VC156 (lane 6), VC157 (lane 7), VC159 (lane 8), VC167 (lane 9), and VC189 (lane 10) all show PCR products of identical size.

Polymorphism of the *flaA* gene of other LIO serogroups. By employing the pg50-RAA19 *flaA* primer pair, we examined the 34 *Campylobacter* strains from 13 other LIO serogroups listed in Table 1. Under the amplification conditions employed, all strains were shown to produce a product indistinguishable in size from the 1,448-bp product from VC167-T2 detected by agarose gel electrophoresis (data not shown). To determine the polymorphic groupings of these strains, the PCR products generated with the pg50-RAA19 primer pair were digested with *Pst*I and *Eco*RI and analyzed on polyacrylamide gels.

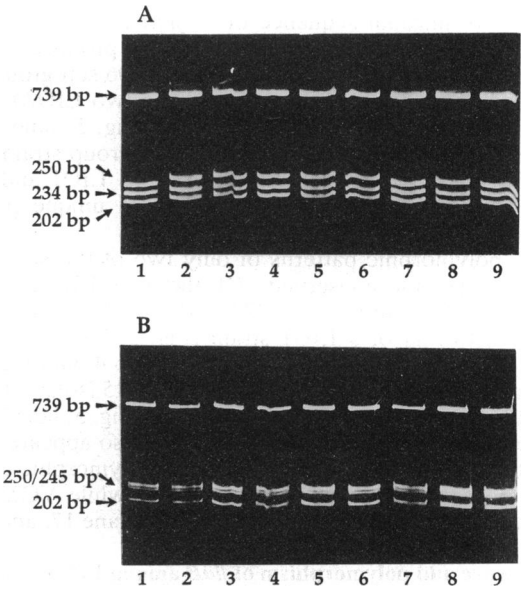

FIG. 4. Restriction fragment length conservation of the PCR fragments seen in Fig. 3 with *Pst*I and *Eco*RI. Strains VC20 (lane 1), VC143 (lane 2), VC144 (lane 3), VC152 (lane 4), VC156 (lane 5), VC157 (lane 6), VC159 (lane 7), VC167 (lane 8), and VC189 (lane 9) all show conserved polymorphic patterns. The fragment sizes are shown on the left.

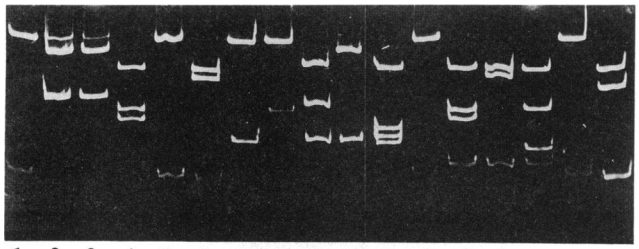

FIG. 5. Restriction fragment length polymorphism generated by *Pst*I and *Eco*RI digestion of the PCR fragments generated with the pg50-RAA19 primer pair from the following strains: VC235 (LIO55, lane 1), VC88 (LIO5, lane 2), 81-116 (LIO6, lane 3), VC83 (LIO4, lane 4), VC94 (LIO15, lane 5), VC87 (LIO1, lane 6), VC104 (LIO19, lane 7), VC92 (LIO12, lane 8), VC95 (LIO7, lane 9), VC97 (LIO20, lane 10), VC167 (LIO8, lane 11), VC74 (LIO11, lane 12), VC230 (LIO11, lane 13), VC232 (LIO11, lane 14), VC91 (LIO11, lane 15), VC168 (LIO29, lane 16), VC236 (LIO29, lane 17). For reference, the fragment sizes in lane 11 are (top to bottom) 739, 250, 232, and 202 bp.

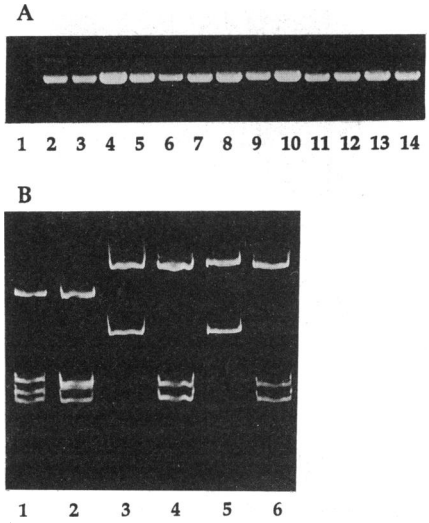

FIG. 6. (A) Products of identical size (1,459 bp) representing *flaB* were identified from all serogroups with the primer pair RAA9-RAA19. Lane 1 is the no-DNA control. The other strains are VC87 (LIO1, lane 2), VC83 (LIO4, lane 3), VC88 (LIO5, lane 4), 81-116 (LIO6, lane 5), VC95 (LIO7, lane 6), VC167 (LIO8, lane 7), VC91 (LIO11, lane 8), VC92 (LIO12, lane 9), VC94 (LIO15, lane 10), VC97 (LIO20, lane 11), VC104 (LIO19, lane 12), VC168 (LIO29, lane 13), and VC235 (LIO55, lane 14). (B) Comparison of polymorphic patterns generated by *Pst*I and *Eco*RI digestion of the *flaA* (lanes 1, 3, and 5) and *flaB* (lanes 2, 4, and 6) genes from LIO8 isolate VC167 (lanes 1 and 2), LIO5 isolate VC88 (lanes 3 and 4), and LIO6 isolate 81-116 (lanes 5 and 6) on a 12% polyacrylamide gel. For reference, the fragment sizes in lane 1 are (top to bottom) 739, 250, 232, and 202 bp.

The restriction profile of a representative strain from each LIO serogroup is shown in Fig. 5. It can be seen that there is a wide range of primary structure polymorphism within the *flaA* genes from the various LIO serotypes on the basis of digestion with these two enzymes, and 10 distinct groupings were identified. In most cases, however, the primary gene structure was well conserved within a given serogroup. All five LIO4 strains possessed the same DNA patterns (Fig. 5, lane 4). The four LIO15 (Fig. 4, lane 5), the three LIO19 (Fig. 5, lane 7), and the two LIO12 (Fig. 5, lane 8) strains were equally conserved within their serogroup (Table 1). In several cases, identical restriction profiles were displayed by strains belonging to two or more different LIO serotypes. The polymorphic patterns of the four LIO6 strains (Fig. 5, lane 3) were identical and matched the expected sizes from the known flagellin sequence of *C. jejuni* 81-116 (27). The LIO5 strains examined were the same and produced a DNA fragment pattern similar to that of the LIO6 serogroup (Fig. 5, lane 2, and Table 1). In addition, the two LIO20 strains (Fig. 5, lane 10) and the LIO55 strain (Fig. 5, lane 1, and Table 1) were identical to the LIO19 serogroup strains (Fig. 5, lane 7). Single isolates from serotypes LIO1 and LIO7 were also examined, and both produced unique patterns (Fig. 4, lanes 6 and 9).

The polymorphic patterns of only two of the serogroups tested were not conserved. Of the five LIO11 isolates examined, two strains, VC228 and VC232, had patterns identical to that of a LIO1 strain (Fig. 5, lane 14); VC230 (Fig. 5, lane 13) was in polymorphic group 4 with the LIO4 strains; VC74 was in group 5 with the LIO15 isolates (Fig. 5, lane 12); and the last, VC91, was unique (Fig. 5, lane 15, and Table 1). Further, the LIO29 serogroup also appeared to be nonconserved, with VC168 falling into polymorphic group 5 with the LIO15 strains (Fig. 5, lane 16) while VC236 produced a unique restriction pattern (Fig. 5, lane 17, and Table 1).

Presence and polymorphism of *flaB* among LIO serogroups. The *flaB* gene has previously been identified with certainty only in *C. coli* VC167 (LIO8) and *C. jejuni* 81-116 (LIO6) (15, 27), although its presence in *C. jejuni* IN1 (LIO7), VC208 (LIO32), VC212 (LIO46), and CL99 (LIO46) has been suggested (12, 43). With the *flaB* primer pair (RAA9-RAA19) that selectively amplified the *flaB* gene in LIO8 isolates, the presence of PCR products of identical size was detected for all serotypes tested, confirming the presence of a second flagellin gene (Fig. 6A). The PCR product generated by amplification with RAA9-RAA19 was also digested with *Pst*I and *Eco*RI. Among the 41 *Campylobacter* strains for which a PCR product was generated with this primer pair, the restriction patterns obtained for 35 isolates belonging to all but two of the serotypes examined were identical to that seen for the corresponding *flaA* gene (summarized in Table 1). The restriction pattern of the *flaB* gene from the four LIO6 strains was the same as that of the *flaB* gene of the two LIO5 isolates examined and was different from the *flaA* patterns of both LIO5 and LIO6 strains (Fig. 6B).

The possibility existed that although the RAA9-RAA19 primer pair can specifically amplify the *flaB* gene product in LIO8, LIO5, and LIO6 strains as determined directly by restriction fragment length polymorphism, these primers could be nonspecifically priming from the *flaA* gene in the other LIO serotypes for which the polymorphic pattern generated by these primers is identical to the pattern generated with the pg50-RAA19 primer pair. To eliminate this possibility, PCR amplification with both primer pairs was performed with the isogenic flagellin mutants from LIO1, LIO4, LIO11, LIO12, LIO15, LIO19, and LIO20 described previously (3). Amplification with the pg50-RAA19 primer pair generated fragments of 2.85 kb from each mutant (data not shown), which corresponded to an increase of 1.4 kb in size due to the presence of the Km^r cassette in the *flaA* gene. Further, amplification with the RAA9-RAA19 primer pair generated wild-type-size fragments of approximately 1.45

TABLE 2. Sequence homologies of *Campylobacter* flagellin genes

Gene	% Homology[a] in region[b]:				
	I	II	III	IV	V
C. coli VC167-T2 *flaA*	100	100	100	100	100
C. coli VC167-T2 *flaB*	87.2	98.4	92.3	100	80.1
C. coli VC167-T1 *flaA*	100	98.4	100	100	100
C. coli VC167-T1 *flaB*	87.2	98.4	94.2	100	80.1
C. jejuni IN1 *flaA*	88.3	80.0	62.8	58.1	87.0
C. jejuni TGH9011 *fla1*	87.4	80.0	70.8	58.5	71.5
C. jejuni 81-116 *flaA*	86.7	86.0	76.6	69.8	88.4
C. jejuni 81-116 *flaB*	82.8	86.4	76.9	69.8	79.7

[a] Compared with the *flaA* gene from *C. coli* VC167-T2 (14).
[b] Regions I to V are as previously defined by Guerry et al. (15).

kb, confirming the presence of a second intact flagellin gene (data not shown).

DISCUSSION

This study has provided further information on the primary structure of *Campylobacter* flagellin genes, especially within the coding areas for the N-terminal and central domains of flagellin. Previous sequencing studies of *Campylobacter* flagellin genes have shown that regardless of the LIO serotype, the regions corresponding to the N and C termini of the flagellin protein are relatively conserved, as indeed they are in the better-described enterobacterial flagellins compared with the central variable domain (Table 2). The one exception in the case of *Campylobacter* flagellin genes sequenced to date is in region V of the *flaA* gene of *C. jejuni* TGH9011, which displays only 71.5% identity to region V of the *flaA* gene of *C. coli* VC167-T2 (14) (Table 2). However, the *fla1* gene encodes a flagellin with a deduced M_r of 56,000, compared with the protein with an M_r of 60,000 produced by *C. coli* VC167. In this study, restriction analysis of specific PCR-generated fragments has demonstrated that the central region of the *flaA* gene is highly polymorphic among strains belonging to various heat-labile LIO serogroups of thermophilic campylobacters. However, in all but two serogroups examined, LIO11 and LIO29, the *flaA* restriction profile is conserved by isolates within a given serogroup.

Using a PCR with oligonucleotide primers shown with LIO8 strains to specifically amplify the *flaB* gene, we have also confirmed the presence of a second flagellin gene among *C. jejuni* and *C. coli* isolates representing 11 serogroups. This is the first study to examine either the distribution or the diversity of *flaB*. For all strains examined except those isolates belonging to the LIO5 and LIO6 serotypes, the polymorphic pattern of the second *fla* gene was identical to that of *flaA*, supporting previous data that demonstrated the high homology of the two flagellin genes (15), and leads us to believe that *C. coli* and *C. jejuni* strains uniformly possess two highly homologous flagellin genes. The reasons for the maintenance of two flagellin genes may be related to regulation by alternate promoters (5) and/or to ensure that motility is maintained (6).

This intragroup conservation of two flagellin genes is the first evidence for the existence of a class of discrete, stable clonal groupings within *Campylobacter* spp., which appear to parallel the LIO serogroups. Although clonality exists in other pathogenic bacteria, with well-studied examples being salmonellae (8, 37), *Neisseria meningitidis* (9, 29), *Escherichia coli* (1, 38, 51, 52), and *Listeria monocytogenes* (34), previous attempts to identify clonal groupings within *Campylobacter* spp. by a variety of immunologic and genetic techniques have been generally unsuccessful. For example, restriction endonuclease analysis, multilocus enzyme electrophoresis, pulsed-field gel electrophoresis, and Southern analysis with rRNA or genes coding for rRNA (ribotyping) have all been employed to examine chromosomal DNA and to attempt to detect minor alterations at the nucleotide level. Indeed, the combination of these sensitive techniques has proven to be useful when attempting to discriminate among strains during epidemiologic evaluations (33). However, the application of individual techniques has provided conflicting evidence with respect to the diversity of *Campylobacter* chromosomal DNA. When multilocus enzyme electrophoresis was used to study the population genetics of *Campylobacter* spp., 50 *C. jejuni* and 14 *C. coli* electrophoretic types were identified with 104 and 21 isolates, respectively, indicating a very high degree of genetic diversity (2). Similarly, by employing ribotyping analysis, Wachsmuth and coworkers (48) demonstrated considerable heterogeneity among outbreak-causing LIO4 isolates, whereas Patton et al. (33) demonstrated reasonably close relationships among separate epidemic-associated isolates of this frequently occurring serogroup.

The conservation of the primary flagellin gene structure within heat-labile serogroups and the polymorphism among serogroups demonstrated here clearly illustrate some form of genetic lineage groups. This finding is somewhat unexpected because the previously observed genetic diversity and apparent lack of clonal groupings suggested that the central domain would be one of the last regions of the flagellin genes to demonstrate conservation. The reason why the flagellin gene structure is so highly conserved within serogroups when the flagellar antigens are rarely the LIO group serodeterminant (3) has yet to be elucidated. However, flagellar antigens do appear to be serospecific and antigenically conserved within a given LIO serogroup (4, 17), a finding which is consistent with the conservation of primary flagellin gene structure within a serogroup. The high level of identity of the *fla* genes within a given strain may be due at least in part to continuous natural transformation and recombination of flagellin information released from neighboring cells. Natural transformation may also be responsible for the duplication of flagellin genes.

Not all of the serogroups examined displayed singular discrete lineage of their flagellin genes. For example, the two LIO29 strains, both of which were distinct from the LIO8/(29) isolates which agglutinate with both the LIO8 and the LIO29 typing antisera, were classed into separate polymorphic groups. The presence of four distinct polymorphic patterns within isolates of the LIO11 serogroup is intriguing. In a previous study, we were unable to isolate an isogenic nonflagellated mutant of VC91 after transformation with VC167B3 genomic DNA in which both the *flaA* and *flaB* genes had been disrupted (3). These data, in conjunction with the nonconservation of the primary structure of LIO11 flagellin genes demonstrated here, indicate that there may be an unusual arrangement of flagellin information in this serogroup and that this serogroup fails to demonstrate clonality. Importantly, however, in all the LIO29 and LIO11 isolates examined, the restriction profile of the *flaB* gene was identical to that of the corresponding *flaA* gene. This configuration of homologous *fla* genes appears to be preferred in *Campylobacter* spp., because the only examples we observed which diverged from this pattern were isolates be-

longing to serogroups LIO5 and LIO6, whose *flaA* and *flaB* genes were in polymorphism groups 2 and 3, respectively (Table 1).

The subunit flagellin size of *Campylobacter* strains is generally large, ranging between 58 and 62 kDa (17), which indicates that the central domain, although shown here to be highly variable at the sequence level and thought to vary antigenically, may be under certain structural constraints. This is further reflected by the conservation in size of the PCR products generated with both the *flaA* and *flaB* primer pairs. This contrasts to the situation seen among serogroups of *Salmonella*, in which the central region often varies in size (40, 49, 50). It has been suggested that the evolutionary mechanism that generates the huge diversity in *Salmonella* flagellum serovars does not occur simply by random mutational drift but by lateral transfer and recombination between preexisting flagellin genes (39). A similar mechanism to generate diversity, especially with two copies of the flagellin gene, may also have occurred among serogroups of *Campylobacter* to create the distinct groups that we observe now.

These data, when considered together with previous findings, provide a significantly expanded model of *Campylobacter* flagella which is consistent with the model for antigenically distinct flagellins of members of the family *Enterobacteriaceae*, in which each *fla* gene codes for a protein with conserved termini and a variable central domain which is conserved within serogroups. Although we have demonstrated that all *Campylobacter* spp. appear to have multiple *fla* genes, the current data do not reveal any information about the organization, regulation, or expression of the two flagellin genes in these other strains. Restriction fragment length polymorphism indicates that in most serogroups both the *flaA* and *flaB* genes are extremely homologous. Within most LIO heat-labile serogroups, there is structural conservation of the two *fla* genes, but among serogroups, the *fla* genes show considerable polymorphism in their central regions. This is reflected at the protein level by flagellins which are antigenically diverse among serogroups but are normally cross-reactive within a given serogroup (4). The possibility does exist that the serospecificity of flagellar epitopes among other heat-labile LIO serogroups is due to posttranslational modifications, as seen with LIO8 (4). These modifications may be independent of the primary gene sequence, whose maintained conservation may be solely an indication of distant clonal groupings of *Campylobacter* isolates.

ACKNOWLEDGMENTS

We thank Corinne Gustafson and Paul O'Toole for valuable discussions.

This work was supported in part by a grant to T.J.T. from the Medical Research Council of Canada and by U.S. Navy Research and Development Command Research Work Unit no. 61102A3M161102BS13 AK.111. R.A.A. was the recipient of an International Postdoctoral Fellowship from the Natural Science and Engineering Research Council of Canada.

REFERENCES

1. **Achtman, M., and G. Pluschke.** 1986. Clonal analysis of descent and virulence among selected *Escherichia coli*. Annu. Rev. Microbiol. **40:**185–210.
2. **Aeschbacher, M., and J. Piffaretti.** 1989. Population genetics of human and enteric *Campylobacter* strains. Infect. Immun. **57:** 1432–1437.
3. **Alm, R. A., P. Guerry, M. E. Power, H. Lior, and T. J. Trust.** 1991. Analysis of the role of flagella in the heat-labile Lior serotyping scheme of thermophilic campylobacters by mutant allele exchange. J. Clin. Microbiol. **29:**2438–2445.
4. **Alm, R. A., P. Guerry, M. E. Power, and T. J. Trust.** 1992. Variation in antigenicity and molecular weight of *Campylobacter coli* VC167 flagellin in different genetic backgrounds. J. Bacteriol. **174:**4230–4238.
5. **Alm, R. A., P. Guerry, and T. J. Trust.** The σ^{54} *flaB* flagellin promoter is subject to environmental regulation. Submitted for publication.
6. **Alm, R. A., P. Guerry, and T. J. Trust.** Significance of duplicated flagellin genes in *Campylobacter*. J. Mol. Biol., in press.
7. **Anderson, P., and J. Roth.** 1981. Spontaneous tandem genetic duplications in *Salmonella typhimurium* arise by unequal recombination between rRNA (*rrn*) cistrons. Proc. Natl. Acad. Sci. USA **78:**3113–3117.
8. **Beltran, P., J. M. Musser, R. Helmuth, J. J. Farmer III, W. M. Frerichs, I. K. Wachsmuth, K. Ferris, A. C. McWhorter, J. G. Wells, A. Cravioto, and R. K. Selander.** 1988. Toward a population genetic analysis of *Salmonella*: genetic diversity and relationships among strains of serotypes *S. choleraesuis*, *S. derby*, *S. dublin*, *S. enteritidis*, *S. heidelberg*, *S. infantis*, *S. newport*, and *S. typhimurium*. Proc. Natl. Acad. Sci. USA **85:**7753–7757.
9. **Caugant, D. A., L. F. Mocca, C. E. Frasch, L. O. Froholm, W. D. Zollinger, and R. K. Selander.** 1987. Genetic structure of *Neisseria meningitidis* populations in relation to serogroup, serotype, and outer membrane protein pattern. J. Bacteriol. **169:**2781–2792.
10. **Ely, B., and L. Shapiro.** 1984. Regulation of cell differentiation in *Caulobacter crescentus*. Microbial development, p. 1–26. *In* R. Losick and L. Shapiro (ed.), Microbial development. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
11. **Fedorov, O. V., A. S. Kostyukova, and M. G. Pyatibratov.** 1988. Architectonics of a bacterial flagellin filament subunit. FEBS Lett. **241:**145–148.
12. **Fischer, S. H., and I. Nachamkin.** 1991. Common and variable domains of the flagellin gene, *flaA*, in *Campylobacter jejuni*. Mol. Microbiol. **5:**1151–1158.
13. **Gerl, L., and M. Sumper.** 1988. Halobacterial flagellins are encoded by a multigene family. J. Biol. Chem. **263:**13246–13251.
14. **Guerry, P., R. A. Alm, M. E. Power, S. M. Logan, and T. J. Trust.** 1991. Role of two flagellin genes in *Campylobacter* motility. J. Bacteriol. **173:**4757–4764.
15. **Guerry, P., R. A. Alm, M. E. Power, and T. J. Trust.** 1992. Molecular and structural analysis of *Campylobacter* flagellin, p. 267–281. *In* I. Nachamkin, M. J. Blaser, and L. Tompkins (ed.), *Campylobacter jejuni*: current status and future trends. American Society for Microbiology, Washington, D.C.
16. **Guerry, P., S. M. Logan, S. Thornton, and T. J. Trust.** 1990. Genomic organization and expression of *Campylobacter* flagellin genes. J. Bacteriol. **172:**1853–1860.
17. **Harris, L. A., S. M. Logan, P. Guerry, and T. J. Trust.** 1987. Antigenic variation of *Campylobacter* flagella. J. Bacteriol. **169:**5066–5071.
18. **Iino, T.** 1977. Genetics of structure and function of bacterial flagella. Annu. Rev. Genet. **11:**161–182.
19. **Joys, T.** 1976. Identification of an antibody binding site in the phase-1 flagellar protein of *Salmonella typhimurium*. Microbios **15:**221–228.
20. **Joys, T.** 1988. The flagellar filament protein. Can. J. Microbiol. **34:**452–458.
21. **Joys, T. M.** 1985. The covalent structure of the phase-1 flagellar filament protein of *Salmonella typhimurium* and its comparisons with other flagellins. J. Biol. Chem. **247:**5180–5193.
22. **Kalmokoff, M. L., and K. F. Jarrell.** 1991. Cloning and sequencing of a multigene family encoding the flagellins of *Methanococcus voltae*. J. Bacteriol. **173:**7113–7125.
23. **Kalmokoff, M. L., S. F. Koval, and K. F. Jarrell.** 1988. Isolation of flagella from the archaebacterium *Methanococcus voltae* by phase separation with Triton X-114. J. Bacteriol. **170:**1752–1758.
24. **Khawaja, R., K. Neote, H. L. Bingham, J. L. Penner, and V. L. Chan.** 1992. Cloning and sequence analysis of the flagellin gene of *Campylobacter jejuni* TGH9011. Curr. Microbiol. **24:**213–221.

25. **Kostrzynska, M., J. D. Betts, J. W. Austin, and T. J. Trust.** 1991. Identification, characterization, and spatial localization of two flagellin species in *Helicobacter pylori* flagella. J. Bacteriol. **173:**937–946.
26. **Kuwajima, G.** 1988. Flagellin domain that affects H antigenicity of *Escherichia coli* K-12. J. Bacteriol. **170:**485–488.
26a. **Lior, H., D. L. Woodward, J. A. Edgar, L. J. Laroche, and P. Gill.** 1982. Serotyping of *Campylobacter jejuni* by slide agglutination based on heat-labile antigenic factors. J. Clin. Microbiol. **15:**761–768.
27. **Nuijten, P. J. M., A. J. A. M. VanAsten, W. Gaastra, and B. A. M. Van der Zeijst.** 1990. Structural and functional analysis of two *Campylobacter jejuni* flagellin genes. J. Biol. Chem. **265:**17798–17804.
28. **Nuijten, P. J. M., B. A. M. van der Zeijst, and D. G. Newell.** 1991. Localization of immunogenic regions on the flagellin proteins of *Campylobacter jejuni* 81116. Infect. Immun. **59:** 1100–1105.
29. **Olyhoek, T., B. A. Crowe, and M. Achtman.** 1987. Clonal population structure of *Neisseria meningitidis* serogroup A isolated from epidemics and pandemics between 1915 and 1983. Rev. Infect. Dis. **9:**665–692.
30. **Oyofo, B. A., S. A. Thornton, D. H. Burr, T. J. Trust, O. R. Pavlovskis, and P. Guerry.** 1992. Specific detection of *Campylobacter jejuni* and *Campylobacter coli* by using polymerase chain reaction. J. Clin. Microbiol. **30:**2613–2619.
31. **Pallesen, L., and P. Hindersson.** 1989. Cloning and sequencing of a *Treponema pallidum* gene encoding a 31.3-kilodalton endoflagellar subunit (FlaB2). Infect. Immun. **57:**2166–2172.
32. **Parales, J., Jr., and E. P. Greenberg.** 1991. N-terminal amino acid sequences and amino acid compositions of the *Spirochaeta aurantia* flagellar filament polypeptides. J. Bacteriol. **173:**1357–1359.
33. **Patton, C. M., I. K. Wachsmuth, G. M. Evins, J. A. Kiehlbauch, B. D. Plikaytis, N. Troup, L. Tompkins, and H. Lior.** 1991. Evaluation of 10 methods to distinguish epidemic-associated *Campylobacter* strains. J. Clin. Microbiol. **29:**680–688.
34. **Piffaretti, J.-C., H. Kressebuch, M. Aeschbacher, J. Bille, E. Bannerman, J. M. Musser, R. K. Selander, and J. Rocourt.** 1989. Genetic characterization of clones of the bacterium *Listeria monocytogenes* causing epidemic disease. Proc. Natl. Acad. Sci. USA **86:**3813–3822.
35. **Pleier, E., and R. Schmitt.** 1989. Identification and sequence analysis of two related flagellin genes in *Rhizobium meliloti*. J. Bacteriol. **171:**1467–1475.
36. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
37. **Selander, R. K., P. Beltran, N. H. Smith, R. M. Barker, P. B. Crichton, D. C. Old, J. M. Musser, and T. S. Whittam.** 1990. Genetic population structure, clonal phylogeny, and pathogenicity of *Salmonella paratyphi* B. Infect. Immun. **58:**1891–1901.
38. **Selander, R. K., D. A. Caugant, and T. S. Whittam.** 1987. Genetic structure and variation in natural populations of *Escherichia coli*, p. 1625–1648. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, vol. 2. American Society for Microbiology, Washington, D.C.
39. **Smith, N. H., P. Beltran, and R. K. Selander.** 1990. Recombination of *Salmonella* phase 1 flagellin genes generates new serovars. J. Bacteriol. **172:**2209–2216.
40. **Smith, N. H., and R. K. Selander.** 1990. Sequence invariance of the antigen-coding central region of the phase 1 flagellar filament gene (*fliC*) among strains of *Salmonella typhimurium*. J. Bacteriol. **172:**603–609.
41. **Southam, G., M. L. Kalmokoff, K. F. Jarrell, S. F. Koval, and T. J. Beveridge.** 1990. Isolation, characterization, and cellular insertion of the flagella from two strains of the archaebacterium *Methanospirillum hungatei*. J. Bacteriol. **172:**3221–3228.
42. **Thomashow, L. S., and S. C. Rittenberg.** 1985. Waveform analysis and structure of flagella and basal complexes from *Bdellovibrio bacteriovorus* 109J. J. Bacteriol. **163:**1038–1046.
43. **Thornton, S. A., S. M. Logan, T. J. Trust, and P. Guerry.** 1990. Polynucleotide sequence relationships among flagellin genes of *Campylobacter jejuni* and *Campylobacter coli*. Infect. Immun. **58:**2686–2689.
44. **Tilsty, D. T., A. M. Albertini, and J. H. Miller.** 1984. Gene amplification in the *lac* region of *E. coli*. Cell **37:**217–224.
45. **Trachtenberg, S., and D. J. DeRosier.** 1988. Three-dimensional reconstruction of the flagellar filament of *Caulobacter crescentus*. A flagellin lacking the outer domain and its amino acid sequence lacking an internal segment. J. Mol. Biol. **202:**787–808.
46. **Vonderviszt, F., S. I. Aizawa, and K. Namba.** 1991. Role of the disordered terminal regions of flagellin in filament formation and stability. J. Mol. Biol. **221:**1461–1474.
47. **Vonderviszt, F., S. Kanto, S. Aizawa, and K. Namba.** 1989. Terminal regions of flagellin are disordered in solution. J. Mol. Biol. **209:**127–133.
48. **Wachsmuth, I. K., J. A. Kiehlbauch, C. A. Bopp, D. N. Cameron, N. A. Stockbine, J. G. Wells, and P. A. Blake.** 1991. The use of plasmid profiles and nucleic acid probes in epidemiologic investigations of foodborne, diarrheal diseases. Int. J. Food Microbiol. **12:**77–90.
49. **Wei, L.-N., and T. M. Joys.** 1985. Covalent structure of three phase-1 flagellar filament proteins of *Salmonella*. J. Mol. Biol. **186:**791–783.
50. **Wei, L.-N., and T. M. Joys.** 1986. The nucleotide sequence of the H-1r gene of *Salmonella rubislaw*. Nucleic Acids Res. **14:**8227.
51. **Whittam, T. S., and R. A. Wilson.** 1988. Genetic relationships among pathogenic *Escherichia coli* of serogroup O157. Infect. Immun. **56:**2467–2473.
52. **Whittam, T. S., and R. A. Wilson.** 1988. Genetic relationships among pathogenic strains of avian *Escherichia coli*. Infect. Immun. **56:**2458–2466.