EPIDEMIOLOGICAL TYPING

Lineages within *Campylobacter jejuni* defined by numerical analysis of pulsed-field gel electrophoretic DNA profiles

JANET GIBSON, E. LORENZ and R. J. OWEN

Laboratory of Enteric Pathogens, Central Public Health Laboratory, 61 Colindale Avenue, London NW9 5HT

Forty-seven Penner heat-stable (HS) serotype reference strains for *Campylobacter jejuni* and 47 serologically non-typable strains were examined by pulsed field gel electrophoresis (PFGE) DNA restriction analysis. The *Sma*I and *Kpn*I digest profiles were compared by numerical analysis. Most strains grouped differently in the two analyses but strain lineages were inferred where the two agreed. Genetic relationships between reference strains in the cross-reacting HS4 complex were examined. Three clonal lines were evident and comprised: (i) HS4, HS13 and HS16; (ii) HS50 and HS65; (iii) HS43. The majority of those *C. jejuni* expressing HS antigens not recognised by currently available antisera had >50% PFGE DNA digest similarity to one or more Penner scheme reference strain(s) and so did not necessarily represent distinct genetic lineages. PFGE analysis provided a high level of discrimination amongst strains of *C. jejuni* but overall similarity estimates for defining types must be based on the analysis of more than one restriction pattern.

Introduction

Examination by pulsed-field gel electrophoresis (PFGE) of genomic DNA restricted with low frequency cutting endonucleases has proved to be a valuable tool in the epidemiological investigation of a wide range of medically important bacteria [1]. The technique has also been applied to the investigation of campylobacter infections including outbreaks involving *Campylobacter jejuni* [2, 3], *C. fetus* [4] and *C. hyointestinalis* [5].

A feature of the epidemiology of human campylobacter infections is that they are mostly sporadic, with relatively few community outbreaks [6]. The sources of the sporadic infections are rarely established, with strains widely distributed in both wild and domesticated birds and mammals, as well as in sewage and untreated water [7]. Because PFGE is a highly discriminatory technique for bacterial strain differentiation, it has considerable potential for tracing the possible routes of transmission of those campylobacters causing the sporadic infections. Comparative studies have established that PFGE profiling can be

highly informative when used in conjunction with conventional typing methods such as biotyping, serotyping and phage typing [8, 9].

The aim of the present study was to apply PFGE combined with numerical analysis to characterise Penner serotype reference strains of *C. jejuni* and recent isolates that were serologically non-typable in the Penner heat-stable (HS) antigen scheme as currently defined. To facilitate the definition of PFGE types, it was necessary to assess criteria for the interpretation of PFGE patterns comprising fewer than 10 bands, and to examine strain groupings obtained with two low frequency cutting enzymes, *Sma*I and *Kpn*I. A previous study showed these to be useful for PFGE analysis of *C. jejuni* [2]. A set of 47 non-typable isolates of *C. jejuni* was examined and compared with the 47 serological reference strains in the *C. jejuni* Penner scheme. The *Sma*I and *Kpn*I banding patterns were analysed by numerical methods to identify matching profiles and to define, where possible, genotypic lineages within *C. jejuni*.

Materials and methods

Bacterial strains and growth conditions

Cultures of reference strains for the 47 different Penner serotypes of *C. jejuni* were obtained from the National

Collection of Type Cultures (NCTC). A further 47 epidemiologically unrelated strains that were non-typable by Penner serotyping were also examined. The strains were received from laboratories in various locations in England and Wales between 1992 and 1994 and included nine strains of animal origin (three chicken, three dog, two lamb, one rat) and 38 isolates from patients with diarrhoeal illness.

All bacteria were cultivated at 37°C for 48 h on defibrinated horse blood 5% v/v agar in micro-aerobic conditions (O₂ 5%, CO₂ 5%, H₂ 2%, N₂ 88%) in a variable atmosphere incubator (Don Whitley Scientific Ltd, Shipley).

Serotyping

Strains were serotyped (heat-stable antigens: HS) according to the passive haemagglutination method of Penner and Hennessy [10] with a panel of 47 antisera provided by Dr A. Lastovica (Red Cross Memorial Childrens Hospital, Cape Town, RSA). The absence of agglutination in a 1 in 80 dilution of antiserum was considered to be a negative reaction.

Pulsed-field gel electrophoresis

The preparation of bacterial DNA for PFGE and the separation of enzyme restriction fragments were performed as described previously [11]. Electrophoretic separation in agarose 1% w/v gels of *Sma*I-generated fragments was at 200 V, 14°C for 22 h with ramped pulse times from 10 s to 35 s; of *Kpn*I generated fragments, 200 V, 14°C for 23 h with ramped pulse times from 4 s to 20 s. DNA degradation in Lior biotype II (DNAase positive) strains was prevented by formaldehyde fixation of the bacterial cells before incorporation into agarose [11].

Analysis of the electrophoretic patterns

Photographic negatives of the PFGE DNA fragment patterns were scanned and the images were captured. Patterns were then compared with GelCompar Software (Applied Maths BVBA, Kortrijk, Belgium). Similarities between the profiles, based on band positions, were derived by the Jaccard correlation coefficient with a maximum position tolerance of 1%. Dendrograms were constructed to reflect the similarities between the strains in the matrix. Strains were clustered by the unweighted pair group method using arithmetic averages (UPGMA). A reference strain DNA digest was loaded in every fourth or fifth well of each gel to facilitate band alignment. The standard DNA marker (λ ladder) was not used for this purpose as the lower bands were not sufficiently sharp for accurate position comparison. The reference strain used for *Sma*I digests was NCTC 12500 and that for *Kpn*I digests was NCTC 12546.

Results

HS serotyping

A total of 1452 human, animal and environmental isolates of *C. jejuni* collected from various geographical locations in the UK over 3 years (1992, 1993 and 1994) were serotyped according to the Penner scheme. A total of 65 isolates (4.5%) was considered to be non-typable with the present panel of 47 typing antisera on the basis that they did not show clearly visible haemagglutination at a serum dilution of greater than 1 in 80. There was no difference between the proportion of non-typable animal and environmental strains (15/322, 4.6%) and the proportion of non-typable human strains (50/1130, 4.4%). The 47 non-typable and epidemiologically unrelated strains were selected for further study from these 65 isolates.

PFGE of *Sma*I and *Kpn*I DNA digests

PFGE of *Sma*I-digested DNA yielded between five and nine fragments ranging in size from 40 to 480 kb (Fig. 1A). The DNA from each of the 47 serologically non-typable strains was cut by *Sma*I although that from the reference strains for HS6 and HS64 was refractory to digestion. *Kpn*I digestion of the DNA samples produced up to 13 easily visualised fragments, most of which ranged in size from c. 40 to 200 kb (Fig. 1B). Smaller bands of low staining intensity could be seen but detection was not reliable and therefore these were not included in the subsequent GelCompar analysis. *Kpn*I digestion produced restriction fragment length polymorphisms (RFLPs) from all but one (C315/94) of the 47 serologically non-typable strains although the DNA from the four reference strains HS22, HS23, HS35 and HS64 was not cut with this enzyme. The DNA from reference strain HS64 was the only DNA of the 94 strains examined not digested by either enzyme.

Numerical analysis of PFGE profiles of reference strains

The dendrogram obtained from numerical analysis of the *Sma*I DNA restriction PFGE profiles of the HS reference strains is shown in Fig. 2. The following strains showed 100% similarity in their *Sma*I fragment patterns: HS9 and HS38; HS42, HS45 and HS57; HS7, HS27 and HS33; and HS50 and HS65. By contrast, four reference strains showed low similarity to all the other strains: HS53 (38%), HS29 (33%), HS43 (32%) and HS63 (17%). Most strains were linked at similarities between 45 and 90%.

The dendrogram obtained from numerical analysis of the *Kpn*I digest fragment patterns for the Penner reference strains (Fig. 3) showed that not all strains grouped in accordance with the *Sma*I-derived analysis. Examples of this were HS9 and HS38 which produced identical *Sma*I patterns, as did HS42, HS45 and HS57, yet there was $\leq 40\%$ band similarity between

Fig. 1. Examples of PFGE profiles of: **A**, *Sma*I and **B**, *Kpn*I digests from strains of *C. jejuni*. The first lane of A and the last lane of B contain size markers (λ ladder, BioRad).

the strains based on analysis of *Kpn*I digests. To highlight the problem of using RFLP patterns from only one enzyme digestion to group strains, the other main anomalies between the two dendrograms are listed in Table 1.

The HS4 cross-reacting serotype complex comprised of HS4, 13, 16, 43 and 50 is well documented [12, 13]. We have extended the group to include HS64 and HS65 as those antisera also reacted with a high proportion of strains that were positive with other antisera in the HS4 complex.

HS4, HS13 and HS16 clustered together in the analyses of profiles produced by both *Sma*I and *Kpn*I. HS13 and HS16 had similar *Sma*I banding patterns with the exception of one band that was ≈ 48 kb larger in HS13 (≈ 218 compared with ≈ 170 kb, respectively) (Fig. 4). The results were consistent with those of the *Kpn*I digest in which HS13 was observed to have ≈ 48 kb of additional DNA and an extra restriction site. With *Sma*I digestion, both HS13 and HS16 showed a 340-kb band that was ≈ 39 kb larger than the 301-kb band present in HS4. Also, the ≈ 218 -kb band in HS13 and HS4 was not present in HS16, although HS16 had a band ≈ 48 kb smaller (≈ 170 kb) that was not seen in the other two strains (Fig. 4).

The *Sma*I generated fragment patterns of HS50 and HS65 were identical, but showed only 23% similarity to those of the HS4, HS13 and HS16 group (Figs 2 and 4). Likewise in the *Kpn*I-generated dendrogram, HS50 and HS65 formed a cluster linked at the 79% similarity level that was well separated from the HS4, 13, 16 group (37% similarity). HS43 was distantly linked to other serotypes of the HS4 complex and

showed only 23% band similarity in *Sma*I digests and 27% in *Kpn*I digests. The DNA of the reference strain HS64 was not cut with either of the enzymes.

Numerical analysis of PFGE profiles for serologically non-typable strains

The banding profiles for the majority of the 47 serologically non-typable strains were $\geq 50\%$ similar to a reference strain on analysis of both *Sma*I and *Kpn*I DNA digests (Table 2; dendrograms not shown). The percentages of strains with $> 40\%$, 60% or 70% *Sma*I or *Kpn*I similarity with a reference pattern are also listed in Table 2. One non-serotypable isolate (C152/94) showed 100% band pattern similarity with HS3 following *Sma*I digestion and 93% with HS3 following *Kpn*I digestion.

Of the five isolates which showed $< 40\%$ similarity to a reference strain with *Sma*I, four formed a cluster with 40% of bands in common (C158/94, C928/94, C351/94, C359/94), but with only 21% similarity to any reference strain. Two of these four strains (C351/94 and C359/94) showed 100% pattern similarity with both enzymes but on the *Kpn*I dendrogram they were 42% similar to a range of serotypes. The fifth isolate (C408/94) was 36% (*Sma*I) and 39% (*Kpn*I) similar to the nearest reference strain (HS40).

When the dendrograms were compared there were many minor variations between the two in the similarities of the non-serotypable strains and the reference strains. Examples where the anomalies were particularly marked are listed in Table 3.

Three pairs of non-serotypable strains showed 100% identity with both enzymes (Table 4). The strains in

Fig. 2. Dendrogram of the cluster analysis of *Sma*I-digested DNA from the Penner (HS) reference strains for *C. jejuni*. HS numbers are shown on the vertical axis. The numbers on the horizontal axis indicate the percentage similarities as determined by the Jaccard correlation coefficient and UPGMA clustering.

each pair were isolated from the same host species; i.e., one pair from dogs, one from humans and one from chickens. Six further pairs were identified following digestion with one of the enzymes but did not show banding identity when the DNA was digested with the other enzyme (Table 4).

Discussion

The sources and routes of transmission of campylobacter enteric infections are not yet clearly established and their elucidation is dependent on the availability of discriminatory typing methods. To date, serotyping

Fig. 3. Dendrogram of the cluster analysis of *KpnI*-digested DNA from the Penner (HS) reference strains for *C. jejuni*. HS numbers are shown on the vertical axis. The numbers on the horizontal axis indicate the percentage similarities as determined by the Jaccard correlation coefficient and UPGMA clustering.

based on either heat-stable (Penner scheme) or heat-labile (Lior scheme) antigens has been used most widely in the characterisation of *C. jejuni*. For both of these typing schemes the proportion of non-typable strains varies from study to study [14]. However, this study demonstrated that only *c.* 5% of the *C. jejuni* isolates randomly collected from various locations in England and Wales over a 3-year period were non-typable with the 47 antisera presently used for the Penner scheme. That figure is similar to data from a comprehensive survey of strains in the USA [15]. Both

studies demonstrated insignificant differences in the proportion of non-typable strains from human and animal or environmental sources. Other estimates of the percentage of non-typable strains of *C. jejuni* from different countries worldwide have varied according to the number of antisera employed, the range of wild animals and birds examined, and the size of the surveys undertaken [14].

The present results support earlier studies [2, 8] in showing that PFGE of DNA digests is highly

Table 1. Comparison of *Sma*I- and *Kpn*I-generated banding patterns for those Penner serotype reference strains of *C. jejuni* that showed the greatest dissimilarity between the two enzyme digests

Serotype reference strains compared	Percentage band similarity*	
	<i>Sma</i> I	<i>Kpn</i> I
HS9 × HS38	100	40
HS42 × HS45 × HS57	100	<40
HS5 × HS15	78	29
HS1 × HS10	81	25
HS10 × HS8	25	77
HS5 × HS9	25	70
HS63 × HS18	17	60

*UPGMA, Jaccard coefficient.

discriminatory for *C. jejuni*. They also show that numerical analysis of such data provides a novel basis for estimating diversity between non-typable and serotype reference strains. Furthermore, the results for the serotype reference strains provide a molecular framework within which the non-typable clinical and animal strains could be assigned an identity. A problem in applying this approach was that current interpretative criteria for PFGE banding patterns, such as those proposed by Tenover *et al.* [1], do not necessarily apply to this study. For instance, the proposed guidelines were not satisfied because the *Sma*I profiles mostly contained < 10 fragments, and the strains, which were collected over an extended period of time, were epidemiologically unrelated. In particular, the analyses demonstrated that estimates of divergence between some strains varied significantly according to the restriction endonuclease used and that matches between PFGE profiles of at least two

enzymes were required to prevent misinterpretation of strain affinities.

The analyses of the PFGE banding patterns of the HS4 complex strains of *C. jejuni* suggest that three serotypes – HS4, HS13 and HS16 – represent a single clonal line. This finding is consistent with that of Preston and Penner [12] who reported only minor differences in the outer-membrane proteins, in restriction endonuclease patterns and in the low mol. wt lipopolysaccharide components of those three serotypes. However, the present study found that the three strains were differentiated by the presence or absence of two DNA fragments of *c.* 39 kb and 48 kb in both enzyme profiles. Such results would be consistent with the integration or loss, or both, of either phage or plasmid DNA by the host chromosome. Antigenic

Table 2. Percentage of non-serotypable strains at various levels of similarity with Penner serotype reference strains of *C. jejuni*

Percentage similarity*	Percentage (number out of 47) of strains	
	<i>Sma</i> I	<i>Kpn</i> I
≥40	89 (42)	94 (44)
≥50	81 (36)	68 (32)
≥60	66 (31)	49 (23)
≥70	45 (21)	15 (7)

*UPGMA, Jaccard coefficient.

Table 3. Comparison of *Sma*I- and *Kpn*I-generated banding pattern similarities between serologically non-typable strains of *C. jejuni* and a Penner reference strain (the examples listed are those where large differences in similarities were observed)

Identity of strains compared	Percentage band similarity*	
	<i>Sma</i> I	<i>Kpn</i> I
C1117/93 × HS5	79	30
C1177/93 × HS17	75	38
C1264/93 × HS8	28	87
C1291/93 × HS17	65	38
C922/94 × HS41	84	27

*UPGMA, Jaccard coefficient.

Table 4. Similarities between PFGE profiles of non-serotypable strains of *C. jejuni* that had identical banding patterns with at least one of the two enzymes used to digest the DNA

Strain (source)	Percentage band similarity*	
	<i>Sma</i> I	<i>Kpn</i> I
C126/93 (H) × C1096/94 (H)	100	100
C140/93 (D) × C143/93 (D)	100	100
C351/94 (C) × C359/94 (C)	100	100
C426/93 (H) × C924/94 (H)	100	62
C688/93 (H) × C824/93 (H)	100	66
C971/93 (H) × C1005/94 (L)	100	60
C216/92 (H) × C967/93 (H)	68	100
C662/94 (H) × C989/94 (H)	100	53
C662/94 (H) × C1001/94 (L)	70	100

H, human; D, dog; C, chicken; L, lamb.

*UPGMA, Jaccard coefficient.

Fig. 4. PFGE of *Sma*I-digested DNA from *C. jejuni* reference strains of the HS4 serotype complex. Lane 1, HS4; 2, HS13; 3, HS16; 4, HS50; 5, HS65; 6, HS43; M, kb marker (lambda ladder).

variation arising from phage conversion has been reported previously in *C. fetus* subsp. *fetus* [16] and is well known in salmonellae. Modulation of O antigens as a result of the introduction of plasmid DNA has been observed in salmonellae [17]. Serotypes HS50 and HS65 also appeared to represent a separate clone, with the two reference strains having identical *Sma*I-generated fragment patterns. Also, serotype HS43 represented a third clonal line within the HS4 complex as it showed < 30% band similarity to other serotypes in that complex. It was not possible to ascertain the clonal identity of HS64 as the DNA from that reference strain could not be cut with either *Sma*I or *Kpn*I. The lack of restriction activity with these enzymes suggests that the strain may possess one or more methylases active in protecting the DNA.

The majority of serologically non-typable isolates of *C. jejuni* showed at least 50% restriction fragment pattern similarity with reference strains and therefore fell within the overall framework of typable strains. One non-typable isolate was closely related to the HS3 reference strain with 100% *Sma*I and 93% *Kpn*I similarity. Although that non-typable strain was genetically similar, it did not appear to express the O side chains on the surface lipopolysaccharides necessary for recognition by the HS3 antisera. Three other serologically non-typable strains were < 40% similar to any reference strain on examination with both enzymes, suggesting that they may represent lineages that were distinct from existing serotypes. However, although a range of different *Sma*I PFGE patterns for the HS1, HS2 and HS4 complex serotypes have already been reported [2, 6], these have not been established for other serotypes. Not all reference strains may be representative of strains within that serotype and it is possible that some of the non-typable strains may be more closely aligned with field strains of an already recognised serotype.

The study demonstrated that PFGE DNA profile analysis is a valuable technique for characterising strains of *C. jejuni* and can be applied to define genotypic relationships between serologically distinct strains as well as between strains that cannot be serotyped at present. Future studies may reveal if other clonal lines in *C. jejuni* have diverged to give rise to more than one serotype or, conversely, if other serotypes have converged to comprise apparently unrelated strain lineages. Such information will be of potential use in the development and improvement of serotyping schemes for the epidemiological investigation of *C. jejuni* infections.

We thank all collaborating laboratories for providing strains. This work was supported by a grant from the Department of Health, London.

References

1. Tenover FC, Arbeit RD, Goering RV *et al.* Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. *J Clin Microbiol* 1995; **33**: 2233–2239.
2. Gibson JR, Fitzgerald C, Owen RJ. Comparison of PFGE, ribotyping and phage-typing in the epidemiological analysis of *Campylobacter jejuni* serotype HS2 infections. *Epidemiol Infect* 1995; **115**: 215–225.
3. Suzuki Y, Ishihara M, Saito M, Ishikawa N, Yokochi T. Discrimination by means of pulsed-field gel electrophoresis between strains of *Campylobacter jejuni* Lior type 4 derived from sporadic cases and from outbreaks of infection. *J Infect* 1994; **29**: 183–187.
4. Rennie RP, Strong D, Taylor DE, Salama SM, Davidson C, Tabor H. *Campylobacter fetus* diarrhea in a Hutterite colony: epidemiological observations and typing of the causative organism. *J Clin Microbiol* 1994; **32**: 721–724.
5. Salama SM, Tabor M, Richter M, Taylor DE. Pulsed-field gel electrophoresis for epidemiologic studies of *Campylobacter hyointestinalis* isolates. *J Clin Microbiol* 1992; **30**: 1982–1984.
6. Healing TD, Greenwood MH, Pearson AD. *Campylobacter* and enteritis. *Rev Med Microbiol* 1992; **3**: 159–167.
7. Cowden J. *Campylobacter*: epidemiological paradoxes. The vehicles for most cases of infection remain unknown. *BMJ* 1992; **305**: 132–133.
8. Owen RJ, Sutherland K, Fitzgerald C, Gibson J, Borman P, Stanley J. Molecular subtyping scheme for serotypes HS1 and HS4 of *Campylobacter jejuni*. *J Clin Microbiol* 1995; **33**: 872–877.
9. Owen RJ, Gibson JR. Update on epidemiological typing of *Campylobacter*. *PHLS Microbiol Dig* 1995; **12**: 2–6.
10. Penner JL, Hennessy JN. Passive hemagglutination technique for serotyping *Campylobacter fetus* subsp. *jejuni* on the basis of soluble heat-stable antigens. *J Clin Microbiol* 1980; **12**: 732–737.
11. Gibson JR, Sutherland K, Owen RJ. Inhibition of DNase activity in PFGE analysis of DNA from *Campylobacter jejuni*. *Lett Appl Microbiol* 1994; **19**: 357–358.
12. Preston MA, Penner JL. Characterization of cross-reacting serotypes of *Campylobacter jejuni*. *Can J Microbiol* 1989; **35**: 265–273.
13. Mills SD, Kuzniar B, Shames B, Kurjanczyk LA, Penner JL. Variation of the O antigen of *Campylobacter jejuni* in vivo. *J Med Microbiol* 1992; **36**: 215–219.
14. Patton CM, Waschmuth K. Typing schemes: are current methods useful? In: Nachamkin I, Blaser MJ, Tompkins LS (eds) *Campylobacter jejuni*: current status and future trends. Washington, DC, American Society for Microbiology. 1992: 110–128.
15. Patton CM, Barrett TJ, Morris GK. Comparison of the Penner and Lior methods for serotyping *Campylobacter* spp. *J Clin Microbiol* 1985; **22**: 558–565.
16. Ogg JE, Chang W. Phage conversion of serotypes in *Vibrio fetus*. *Am J Vet Res* 1972; **33**: 1023–1029.
17. Le Minor L, Rohde R, Charle-Marsaines C, Coynault C. [Antigenic similarities of certain *Salmonella* O:5,4 serotypes with other O groups of the Kaufmann-White scheme.] Nouvelles observations sur les parentés antigéniques de certains sérotypes de *Salmonella* du groupe O:5,4 avec d'autres groupes O du schéma de Kauffmann-White. *Ann Microbiol (Paris)* 1973; **124**: 451–461.