Identification of outbreak-associated and other strains of
Clostridium difficile by numerical analysis of SDS–PAGE protein
patterns

M. COSTAS¹, B. HOLMES¹, M. GANNER¹, S. L. W. ON¹, P. N. HOFFMAN², M. A. WORSLEY⁴ AND H. PANIGRAHI³

¹National Collection of Type Cultures and ²Laboratory of Hospital Infection, Central Public Health Laboratory, London NW9 5HT
³Department of Microbiology and ⁴Infection Control, North Manchester General Hospital, Crumpsall, Manchester M8 6RB, England

(Accepted 21 March 1994)

SUMMARY

Seventy-three cultures of Clostridium difficile isolated both during, and in the period immediately following, an outbreak of infection in a group of three hospitals, were characterized by one-dimensional sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS–PAGE) of whole-cell proteins. Each protein pattern was characterized by the presence of one or two dense bands which were highly reproducible. The protein patterns were used as the basis for a numerical analysis which divided the strains into five phenons (electrophoretic or EP types). The majority, 60 of the 73 cultures, belonged to a single phenon which included strains from both patients and the environment. We conclude that high-resolution SDS–PAGE of proteins provides an effective method for typing C. difficile and therefore for tracing the possible spread of epidemic strains in hospitals and other institutions, thereby allowing a better understanding of the epidemiology of the organism.

INTRODUCTION

The most serious form of disease caused by Clostridium difficile is pseudomembranous colitis (PMC) and in such cases the pathogenic role of the organism is usually clearly apparent [1]. However, the classic symptoms of this illness may be absent and the role of C. difficile in infection may be difficult to determine, especially where asymptomatic colonization is commonly seen [2]. Between these two extremes, the organism is widely implicated in adult hospitalized patients as the major cause of diarrhoea associated with antimicrobial (AAD) and/or antineoplastic chemotherapy [3–5].

Strains of C. difficile produce at least two toxins: enterotoxin (toxin A) and cytotoxin (toxin B). Although these may provide the organism with much of its pathogenic potential, the spectrum and severity of clinical presentations associated with the organism may be because of variance between strains in their
intrinsic ability to cause disease. A number of clinical studies also support the possibility of strain-specific differences in virulence [6-9].

Several factors indicate that intestinal disease induced by *C. difficile* is frequently nosocomial in origin. Firstly, it is apparent that treatment with certain antibiotics increases susceptibility to infection by altering or displacing the balance of the normal bowel flora [10]. Secondly, there is evidence for person-to-person transmission of *C. difficile* in hospitals [11, 12]. Thirdly, there are documented cases of contamination of the hands of health-care workers as well as of environmental contamination [13, 14]. Within the hospital, the mode of transmission of *C. difficile* has, understandably, been generally evaluated in epidemic settings [7, 12, 15, 16].

For epidemiological studies, a variety of methods have been applied to the typing of *C. difficile*: bacteriophage and bacteriocin typing [11, 17, 18], serotyping [19-21], protein typing (including radio-labelled [12, 16], EDTA-extracted surface [7, 22], and whole-cell proteins [23-25]), immunoblotting techniques [26-28], plasmid profiling [29], restriction endonuclease fingerprinting [30, 31], restriction fragment length polymorphisms [32], the polymerase chain reaction [33], and pyrolysis mass spectrometry [34].

The aim of the present study was to compare the high-resolution 1-D SDS-PAGE whole-cell protein patterns of a number of cultures isolated during an outbreak of *C. difficile* infection which took place in three North Manchester hospitals (North Manchester General, Monsall and Northern) between 1 November 1991 and 1 May 1992. During this outbreak, which is fully described elsewhere [35], infection contributed to 17 deaths amongst the mainly geriatric patients. The investigation described herein was prompted by the high isolation rate and spread of *C. difficile* within this hospital group and by the high mortality associated with the outbreak. A computerized analysis of protein patterns [36] was employed to evaluate their use as a method of differentiating the strains involved and further characterizing the epidemiology of the outbreak. Although the study was not prospective, it was fortunate that some strains had been preserved, for comparison, from an earlier outbreak [34, 37].

**MATERIALS AND METHODS**

**Sources and strains**

A total of 79 cultures of *Clostridium difficile* were examined. Of these, most (58) were recovered from 54 patients (4 second specimens after relapse) with either AAD or PMC whilst the remaining 15 were environmental samples. The isolates selected represent a convenience sample set from the 175 patients who developed diarrhoea associated with *C. difficile*. Most of the isolations were made at North Manchester General Hospital (137/175 patients) where the outbreak was thought to have originated. Three index cases were followed 2 weeks later by a further 12 cases in the same acute geriatric ward. Since then the ‘epidemic’ strain has been recovered from a total of 34 wards over three hospitals and appears to have become endemic in North Manchester General Hospital, at least. Eleven of the isolates, all from patients, were isolated in late 1992 after the defined period of the outbreak. In addition, four isolates from a previous outbreak at the same hospitals
Protein patterns of Clostridium difficile in February to April 1991 [34, 37] were included for comparison, together with a reference strain of each of EP types 1 and 12 [38].

The case definition for C. difficile diarrhoea and the patient sampling policy depended on a number of criteria. Isolation of C. difficile was generally only attempted when these criteria were met: (a) patients presented with explosive, watery diarrhoea; (b) hospital in-patients who had been on antibiotics known to predispose to this condition within the past 6 weeks (e.g. ampicillin and cephalosporins); (c) either cytotoxin and/or enterotoxin were detected in faeces. It was considered that C. difficile had been the causative agent of the diarrhoea if, in addition to the above criteria, it was the only enteric pathogen found.

Isolation, culture media and conventional biochemical tests

C. difficile strains were isolated from faeces by inoculation onto selective CCFA medium, comprising C. difficile Agar Base (Oxoid, Unipath Ltd, Basingstoke, Hants, UK) 6-9% (wt/vol), pH 7-4, with horse blood 6-6% (vol/vol), and containing cycloserine (500 μg/ml), cefoxitin (15 μg/ml) and fructose 0-6% (wt/vol). The primary isolation plates were incubated for 48 h at 35 °C in an anaerobic jar with an Oxoid Gas Generating Pack (84% H₂ and 16% CO₂). C. difficile strains were presumptively identified by their characteristic ground-glass appearance on CCFA medium, by their odour, Gram stain and by biochemical tests using the API ATB 32 A kit (bioMérieux UK Ltd, Basingstoke, Hants, UK). Identification as C. difficile was confirmed by a latex agglutination test (Microscreen; Mercia Diagnostics Ltd, Guildford, Surrey, UK).

Strains were suspended in 10% (vol/vol) glycerol broth and maintained under liquid nitrogen. For production of protein samples, cultures were grown on Fastidious Anaerobe Agar (LAB 90 FAA; Lab M, Bury, England) supplemented with 5% (vol/vol) defibrinated horse blood for 24 h at 37 °C in an anaerobic cabinet (Don Whitley Scientific Ltd, Shipley, Yorks, England).

Toxin assay

Analysis of faeces for the presence of C. difficile toxin A was by an ELISA kit (Premier; Meridian Diagnostics Europe, Srl., Milan, Italy). Testing for toxin B was performed by tissue-culture cell assay using MRC V cells (human embryonic lung fibroblast) by modification of a method described previously [39].

Preparation of protein samples, electrophoresis, staining and scanning of gels

Bacterial samples (10 mg wet weight) were harvested directly from FAA plates into 150 μl of lysozyme (3 mg/ml) and incubated for 2 h at 37 °C. An equal volume of double-strength lysis buffer (20% vol/vol glycerol, 2% vol/vol 2-mercaptoethanol, 4% wt/vol SDS and 70% vol/vol stacking gel buffer) was then added and the samples heated at 100 °C for 10 min in a heating block. The protein samples were then extracted as described previously [36].

Samples were run on discontinuous SDS–polyacrylamide gels which were cast to allow for a 10 mm stacking gel. The final polyacrylamide concentrations were 10% wt/vol for the separation gel and 5% wt/vol for the stacking gel. Full details of the methods used in gel preparation and electrophoresis have been described previously [36].
The stained protein patterns in the dried gels were scanned with the equipment and methods described by Costas [36]. The absorbance range was set from 0.1 to 1.2 absorbance units.

**Analysis and computation of similarity**

Analysis and computation of similarity were as described by Costas and colleagues [40] except that the protein patterns were corrected for gel-to-gel variation by segmented linear correction using 17 discernible marker positions on the reference pattern (that of Past 27, a strain of Pasteurella sp. chosen because of its evenly distributed protein band pattern) and linear correction was carried out within each of the 16 defined segments for each track on the calibrated gel. The length-corrected traces on the reference gel were composed of 432 absorbance values (after removal of the initial and final bands) and the background cut-off was set at 0.4. The best fit between each pair of traces was obtained by laterally shifting one corrected trace with respect to the other in single point steps of 160 μm up to three points on either side of the initial alignment. The analysis was based on the whole of the protein pattern and included proteins to 31.5 kDa.

A database was created with the patterns of 58 strains of *C. difficile*. These represented the full diversity of protein pattern types observed in the strains of the present study. The strains (CD203/92–CD318/92) were those used in the cluster analysis (Fig. 2). The remaining 19 strains were allocated to a particular EP type using an identification program. Utilizing the Pearson product moment correlation coefficient, the program calculated and ordered those five strains in the database having the highest similarity to the unknown. Final placement of an unknown strain to a particular EP type was determined by the highest similarity achieved and subsequently confirmed by introduction into the database and UPGMA clustering.

**RESULTS**

**Study population**

All of the study population (*n* = 175) was, by definition, symptomatic in that individuals fulfilled the selection criteria detailed earlier. Although the age range of the patients was from 26 to 101 years, the majority of the population were elderly with 90% of cases being reported from patients of ≥ 60 years of age. The subpopulation whose isolates were analysed here did not differ from the outbreak population as a whole with regard to age and sex distribution. The male-to-female ratio was 42:58. There was a wide range of presenting conditions prior to the onset of diarrhoea, principally pneumonia and cerebrovascular accident. The antibiotics being administered to the patients (169/175) comprised primarily cephalosporins and β-lactamases. All strains were found to be toxin positive.

**Typing of *C. difficile***

One-dimensional SDS–PAGE of whole-cell protein extracts of the 79 isolates included in this study produced patterns containing approximately 40 discrete bands with molecular weights of 18–100 kDa. Proteins of < 18 kDa were not resolved under the electrophoretic conditions used in this study. Most of the bands were relatively weakly stained but there were several prominent bands in the
Protein patterns of Clostridium difficile

Fig. 1. Representative electrophoretic protein patterns of Clostridium difficile EP types. Track 1. EP type 1 reference strain; 2. EP type 1 (Manchester); 3. EP type 12 reference strain; 4. EP type 12 (Manchester); 5. EP type 19; 6 and 7. EP type 20; 8. EP type 18. Molecular weight markers (track labelled X) are (from top to bottom): ovotransferrin, 76-78 kDa; albumin 66.25 kDa; ovalbumin, 42.7 kDa; carbonic anhydrase 30 kDa; myoglobin, 17.2 kDa; R, reference pattern (Past 27).

Numerical analysis of PAGE total protein profiles based on the determination of the Pearson product moment correlation coefficient and UPGMA clustering revealed that, at the 86% (S) similarity level, the 58 C. difficile isolates formed five distinct phenons, as shown in the dendrogram (Fig. 2). Two of the phenons corresponded to types that had been recognized in a previous study (EP types 1 and 12 [38]), but three represented new types, i.e. EP types 18, 19 and 20. The region between 33 and 55 kDa; a number of these were common to all strains and were therefore species-specific. However, within this region, the mobility of one or two prominent bands varied between strains. This served both to group strains where common band patterns were evident and to differentiate these groups on the basis of pattern heterogeneity. Representative PAGE protein patterns for each EP type recognized in this study together with patterns of reference EP types [38], are illustrated in Figure 1. The protein patterns of the isolates examined were highly reproducible both within and between gels. Replicate protein samples of Past 27 and the molecular weight standards run on each gel gave similarity values of 95.9 ± 1.7% and 96.1 ± 1.2%, respectively.
Fig. 2. Dendrogram of the cluster analysis based on total protein content of the 58 *C. difficile* strains from patients involved in the major outbreak (vertical axis). The numbers on the horizontal axis indicate the percentage similarities as determined by the Pearson product moment correlation coefficient and UPGMA clustering. Phenons were formed at the 86% similarity level.
Protein patterns of Clostridium difficile

remaining 19 strains were assigned to a particular type after comparing their individual patterns against the database. Identity was only accepted if the five most similar strains selected were of a single type and they gave similarities well above the cut-off threshold for phenon formation.

Primary patient isolates and environmental isolates

In the primary isolates from patients, 45/54 (83%) were found to be EP type 1, the predominant outbreak strain. Only 3/54 (5.5%) were EP type 12 and EP types 18–20 were each found in only 2/54 (3.7% each).

The 15 environmental isolates were recovered from several wards from two of the hospitals. All eight from the nominated isolation ward for C. difficile cases were found to be the outbreak type, EP type 1; three of these were associated with a particular patient from whom this type was also isolated. Of the remaining 7 isolates, 3 were of EP type 12 and had been isolated from 2 other wards on the same hospital; the 4 remaining strains were all EP type 1, 3 of which were from another hospital.

Over all the isolates, EP type 1 was by far the most prevalent (64/77), followed by EP type 12 (7/77). Together these types accounted for 92% (71/77) of all isolates examined.

Multiple isolates from single patients

There were four patients from whom the organism was isolated on a second occasion (i.e. after relapse). Each pair yielded the same type (three EP 1 and one EP 12) on both occasions the organism was isolated, despite sampling on different wards in the case of two of these patients.

Distribution of EP type 1 by hospital and ward

There were 34 wards over the three hospitals thought to be involved in the outbreak. Cultures isolated from 20 of these wards were included in the present study and 18 of these yielded EP type 1 strains. Strains of EP type 1 predominated in all three hospitals, North Manchester General: 84%, Monsall: 82% and Northern: 89%. A single strain from Ancoats Hospital (an orthopaedic hospital in the same district), which was not thought to be part of the outbreak, proved to be EP type 12. Few isolations were made on surgical wards (5/6 were EP type 1) with the majority of isolations from geriatric wards, particularly the ward which became the designated isolation ward for those harbouring the organism. There were two community-acquired cases, one patient harboured EP type 1 and the other EP type 18. The four isolates from the earlier outbreak [34. 37] also proved to be EP type 1.

DISCUSSION

In the main, C. difficile diarrhoea appears to be an infection of the elderly, where it is associated with prior antimicrobial therapy and toxin production by the organism.

This study was not prospective but included isolates from both hospitalized symptomatic cases and their associated environment. The strains examined did
not represent all the cases of *C. difficile* diarrhoea found in the group of three hospitals over a defined period. However, it did include 4 isolates from an earlier outbreak [34, 37], 62 cultures from patients and the environment which were a randomly selected set from 175 patients who developed diarrhoea associated with *C. difficile* during a subsequent major outbreak [35] and 11 patient isolates following the defined period of the major outbreak. Since EP type 1 was found during all three time periods and was the predominant type through the outbreak described here, it is clear that this type has been present in the hospital group for a considerable period. During this time it appears to have been responsible for two outbreaks of infection, and our findings suggest that it has subsequently become endemic.

The protein electrophoretic typing method used here was previously applied to the analysis of strains from a Canadian hospital which included an outbreak [38]. In that study, EP type 1 predominated and was responsible for the outbreak. It is thus of interest to note that the same EP type has been responsible for outbreaks of infection on two continents. This is clearly an area for further study but it does lend support to the suggestions of others that there may be strain-specific differences in virulence [6–9] and that specific types are associated with major outbreaks [7, 12, 16]. In the Canadian study, more electrophoretypes (17) were found to be present in the hospital than in the Manchester group (5) and EP type 1 accounted for a lower proportion of the total isolates examined in the former. However, the Manchester isolates were selected to represent outbreaks and not sporadic cases. There was, therefore, a bias in selection against background incidence in the Manchester group.

Only two of the Manchester patients had AAD and harboured the organism on admission during the major outbreak. The remaining 52 patients developed diarrhoea while in hospital. The predominance of a single EP type may indicate a common exogenous source for acquisition of the outbreak strain. This is further supported by the finding that the majority of environmental isolates were also of EP type 1. However, this is not the exclusive method of acquiring infection since a previous study [38] indicated that as well as being an exogenous disease, *C. difficile*-associated diarrhoea can also be an endogenous disease (which implies overgrowth of the organism in a susceptible host). The prospective study of McFarland and colleagues [27] showed that 13 of 17 types were isolated from asymptomatic patients on admission to hospital but that many of these same types were subsequently recovered from symptomatic cases. The implication is that many cases of AAD associated with *C. difficile* are imported to the hospital endogenously by the patients. At least some of the Manchester patients, who harboured strains of EP types 12 and 18–20, may have been endogenous cases although EP type 12 strains were also isolated from the environment and may represent a second exogenous source.

The immediate environment of infected patients has been found by a number of workers [41, 42] to be contaminated with *C. difficile*, and in some studies, typing has confirmed that environmental isolates are often of the same type as those of the infected patient with which they are associated [14, 28]. The Manchester environmental samples included two samples from bed linen and the floor associated with one patient, plus a cubicle curtain sample from a second patient:
all were of the same EP type as the patient isolates themselves, i.e. EP type 1. It appears likely that most environmental isolates have been shed by infected patients and consequently the environment can act as a reservoir for further infection by either direct transfer or indirect transmission via a carrier. This is further supported by our findings that the predominant types in the patients, EP types 1 and 12, were the only types found amongst the environmental samples.

Patients can relapse following apparently successful treatment of their diarrhoea. Relapse may be due either to strains of a different type or the same strain may again be implicated [15, 32, 38, 43, 44]. In the main Manchester outbreak there was an overall relapse rate of 12.5% [35]. Samples from only four such patients were included in this study and these were probable instances of recrudescence as they each had the same type on both occasions. EP type 1 accounted for three cases but since this was the outbreak type, re-infection may have been a distinct possibility. The fourth case was an EP type 12 and since this type was the second most common overall, there is again the possibility of re-infection.

Several methods have been applied to the typing of \textit{C. difficile} with varying degrees of success. Established methods, such as bacteriophage and bacteriocin typing, can provide good discrimination but both typability and reproducibility are often poor. Such methods are labour intensive and interpretation can be subjective. Despite being available for several years, no standardized reference set has been established resulting in limited portability. Serological methods are inexpensive and rapid and have evolved to give from 15 to 20 serogroups. Although typability and discrimination are good, it is known that cross-reactions occur. At least one of the major serogroups (A) can be further divided into at least 12 protein profiles. Although useful in specific cases, plasmid profiles cannot act as a universal system of typing because only a small proportion of strains harbour plasmids. Restriction endonuclease (REA) fingerprinting, restriction fragment length polymorphisms and methods based on the polymerase chain reaction are technically complex and expensive. They are generally used in fingerprinting of strains and not for typing. However, recently a typing system based on REA patterns has been described [45] from an examination of almost 2000 strains from which 206 unique types were recognized. The banding patterns though, were extremely complex and so visually difficult to interpret. Such a scheme suffers a severe practical disadvantage unless it is coupled with a sophisticated data capture and computer analysis system. Pyrolysis mass spectrometry has been applied in the investigation of an outbreak but its discriminatory ability has not so far been examined.

Using the SDS–PAGE technique, 17 protein electrophoretypes (EP types 1–17) were defined in a previous study [38]. Three additional EP types (18–20) have been recognized in this study. Although these studies were confined to only two hospital outbreaks, 20 EP types have now been described, similar to the number of types recognized by other methods. Thus, SDS–PAGE is useful for typing strains of \textit{C. difficile} and appears to offer a better, or similar, degree of discrimination to most of the other typing schemes. Differing electrophoretic methodologies have been employed in the typing of \textit{C. difficile} by protein patterns and it appears to be the most popular general approach. These include radio-
M. COSTAS AND OTHERS

labelled [12, 16], EDTA-extracted surface proteins [7, 22], and whole-cell proteins [23–25] and, in addition, these may be coupled with immunoblotting techniques [26–28]. Discrimination is extremely high and in theory all strains can be typed; in practice, well over 95% of isolates can be typed to previously defined types and any new patterns can be added as new types. Discrimination varies according to technique and the method used to define types. The radio-labelled technique has low discrimination with only nine types currently defined compared with more than 17 types using whole-cell or EDTA-extracted protein patterns. The radio-labelled protein method is more expensive and requires special facilities whereas EDTA-extraction and immunoblotting involve additional procedures but it is debatable whether they produce any improvement in discrimination.

In conclusion, whole-cell protein typing by SDS-PAGE can be successfully applied to the epidemiological investigation of \textit{C. difficile} outbreaks. It is relatively cheap, is not technically complex, all strains are typable and the same approach can be used for other organisms. The use of a high-resolution scanner coupled with sophisticated computer software for the analysis of patterns, although not obligatory, allows for the objective comparison of large sets of patterns produced on different gels.

ACKNOWLEDGEMENTS

We thank the Department of Medical Illustration at the Central Public Health Laboratory for photographic assistance.

REFERENCES

Protein patterns of Clostridium difficile


35. Cartmill TDI, Panigrahi H, Worsley MA, McCann DC, Nice CN, Keith E. Management and


