

MOLECULAR EPIDEMIOLOGY

# Clonal groups of enteropathogenic *Escherichia coli* isolated in case-control studies of diarrhoea in Bangladesh

M. ANSARUZZAMAN, M. J. ALBERT, S. NAHAR, R. BYUN, M. KATOULI\*, I. KÜHN\* and R. MÖLLBY\*

International Centre for Diarrhoeal Disease Research, Bangladesh, Dhaka, Bangladesh and \*Microbiology and Tumorbiology Centre, Karolinska Institute, Stockholm, Sweden

Recent case-control studies in Bangladesh showed a high prevalence of enteropathogenic *Escherichia coli* (EPEC) strains (identified by DNA probes for virulence genes) associated with childhood diarrhoea. However, the clonal status of these strains is not known. A total of 94 EPEC isolates from 80 children with diarrhoea and 14 healthy matched controls isolated during 1991–1992 and 1993–1994 was characterised by serogrouping, enterobacterial repetitive intergenic consensus sequence PCR, and by a biochemical fingerprinting method (the phene plate or PhP system). Twelve O serogroups were found with O114 (n = 19) and O127 (n = 23) being the dominant serogroups. Most strains of O114 belonged to the same PhP/PCR types. Strains of O127 contained 16 that produced cytolethal distending toxin (CDT) and seven that did not; both were found among patients as well as controls. Results of PCR and PhP typing showed that CDT-positive strains belonged to the same clonal group and were related to one of the two PhP/PCR types of CDT-negative O127 strains. Thirty-one EPEC strains were O non-typable and 21 strains belonged to other less prevalent serogroups. These strains belonged to diverse PhP/PCR types and did not show any similarity to the strains of two major serogroups, O114 and O127. The results suggest that two clonal groups of EPEC strains are predominantly associated with childhood diarrhoea in Bangladesh.

## Introduction

Diarrhoea is one of the most common causes of morbidity and mortality among infants and young children in developing countries, with *Escherichia coli* being one of the major causative agents [1–3]. Intensive studies have shown the presence of several classes of diarrhoea-producing *E. coli*, each characterised by a number of pathogenic properties. These include enteropathogenic *E. coli* (EPEC) [4–7], enterotoxigenic *E. coli* (ETEC) [8], enteroinvasive *E. coli* (EIEC) [9], enterohaemorrhagic *E. coli* (EHEC) [10], diffuse adherent *E. coli* (DAEC) and enteroaggregative *E. coli* (EAaggEC) [9, 11]. The prevalence of each of these types varies between countries [5, 12, 13].

Although EPEC was the first category of diarrhoea-producing *E. coli* to be identified in the early part of this century, their pathogenic mechanisms have been elucidated only recently. The classical method for identification of EPEC strains is serotyping. This method has served as the main diagnostic tool for many years and up to 12 serogroups have been recognised as the traditional EPEC serogroups [11]. However, the development of DNA probes and cell culture assays has reduced the dependence on serotyping alone for identification of EPEC strains [4, 14, 15]. With these newly developed techniques and in combination with O-serogrouping, many workers have investigated the prevalence of EPEC strains among infants and young children with diarrhoea in different countries [8, 16].

Received 9 Feb. 1999; revised version received 22 June 1999; accepted 11 July 1999.

Corresponding author: Dr M. Katouli (e-mail: mkatouli@usc.edu.au). Present address: Faculty of Science, University of the Sunshine Coast, Maroochydore DC, Queensland 4558, Australia.

A new toxin, named cytolethal-distending toxin (CDT) has been identified among several enteropathogenic bacteria, including EPEC strains [17]. In two recent studies [16, 18] on the aetiology of childhood diarrhoea in Bangladesh, EPEC have been identified as a leading

cause of bacterial diarrhoea in infants, with some strains also producing CDT [18]. In view of the above findings, the present study was done to investigate the clonal status of EPEC strains isolated from these case-control studies in Bangladesh.

## Materials and methods

### EPEC isolates

The EPEC isolates were from two studies conducted at the Dhaka Treatment Centre of the International Centre for Diarrhoeal Disease Research, Bangladesh (ICDDR, B). In the first study, conducted during July 1991–May 1992 [16], the prevalence of diarrhoeagenic *E. coli*, including EPEC, was investigated in 451 children up to 5 years of age with diarrhoea, and in 602 matched control children without diarrhoea. EPEC were identified by hybridisation of *E. coli* colonies with DNA probes for EPEC adherence factor (EAF) and attaching and effacing (AE) factor. Colonies that were hybridisation-positive with both probes or with the AE probe only were considered to be EPEC. In that study, EPEC strains were isolated from 70 (15.3%) patients with diarrhoea and 33 (5.3%) controls, thus showing a significant association ( $p < 0.0001$ ) with diarrhoea.

The second study consisted of 546 children up to 5 years of age, with diarrhoea, attending the same centre at the ICDDR, B during July 1993–May 1994. The control group consisted of 215 matched healthy children. In that study, EPEC were identified by DNA probe hybridisation assay as in the first study. In addition, cytolethal-distending toxin (CDT)-producing *E. coli* were identified by a DNA probe hybridisation assay and tissue culture assay. CDT<sup>+</sup> *E. coli* were isolated from 17 (3.1%) patients and two (0.93%) controls ( $p = 0.082$ ). The isolates from 13 patients and the controls were available for serotyping, and all of them belonged to the serotype O127:H<sup>-</sup> [18]. All the 15 isolates from the second study and the seven O127 isolates from the first study were included in the present study. A summary of the isolates tested in the present study is shown in Table 1.

### Enterobacterial repetitive intergenic consensus (ERIC) PCR

Bacterial DNA was prepared by directly picking colonies from MacConkey agar plates and lysing them in de-ionised water by boiling for 5 min. The supernate was used as the template DNA for PCR. Amplification was performed with a DNA thermal cycler (Perkin-Elmer DNA thermal cyclers, Norwalk, CT, USA) under the following conditions: 25 cycles of 30 s at 94°C, 30 s at 36°C and 90 s at 72°C with 3 min initial denaturation at 94°C and a 10-min final extension at 72°C, with a 50- $\mu$ l volume containing 5  $\mu$ l of bacterial DNA and 0.75 unit of Taq DNA polymerase (Gibco-BRL) in a buffer containing 25 mM MgCl<sub>2</sub>, 200 mM

Table 1. Enteropathogenic *E. coli* (EPEC) isolates included in the present study from the previous two case-control studies of diarrhoea in Bangladesh

Study no.	Serogroup	Number of isolates from		Total number of isolates
		patients	controls	
1	O127	8	0	8
	O114	19	0	19
	O142	2	0	2
	O128	1	0	1
	O126	3	0	3
	O125	3	2	5
	O119	2	0	2
	O86	1	0	1
	O44	1	0	1
	O28	1	0	1
	O20	2	2	4
	O2	1	0	1
	ONT	23	8	31
	O127:H <sup>-</sup>	13	2	15
Total	80	14	94	

EPEC were identified by DNA probe hybridisation. The first study was conducted during July 1991–May 1992, on 451 patients and 602 controls [16]; the second during July 1993–May 1994 on 546 patients and 215 controls [18]. From the second study only CDT-positive EPEC were included. ONT, O antigen non-typable.

Tris-HCl (pH 8.4), 500 mM KCl, 10 mM each of dATP, dCTP, dGTP and dTTP, and 0.25 mM primer. The primer was based on the enterobacterial repetitive consensus sequence (5'-ATGTAAGCTCCTGGGATT-CAC-3') [19]. After PCR amplification, samples of the products were electrophoresed in agarose 1% gel containing ethidium bromide 0.2 mg/ml and photographed under UV light. The gel-photos were scanned by a hand scanner and the reconstructed gel images were created and analysed with GelCompare software, version 4.0 (Applied Maths, Kortrijk, Belgium). DNA fragment patterns generated by PCR were compared pairwise. Isolates with indistinguishable bands were regarded as identical and assigned to a common PCR type. These types were denoted by low-case letters (i.e., a, b, c, etc.). On the basis of differences in size and number of the fragments, isolates with PCR patterns similar to the common types were considered to be subtypes and were indicated by a number after the PCR-designated letter (i.e., subtype of PCR type a is a1). Isolates with PCR patterns distinctly different (< 50% fragments in common) from other types were regarded as non-related or single (si) PCR types [20].

### Biochemical fingerprinting

*E. coli* strains were typed on the basis of their biochemical phenotypes by the Phene Plate (PhP) system specifically developed for typing *E. coli* strains (PhPlate Stockholm AB, Sweden). Briefly, a loopful of a fresh culture of bacteria was suspended in 8 ml of sterile proteose-peptone 0.1% w/v solution containing bromothymol blue 0.11% w/v and 150- $\mu$ l volumes of the suspensions were inoculated into 24 wells in pre-prepared microtitration plates. Plates were stored at 4°C overnight and incubated at 37°C the following morning.

The  $A_{620}$  of each reaction was measured at 7, 24 and 48 h with a microplate reader and the data were automatically transferred to a personal computer, multiplied by 10, yielding scores ranging from 0 to 30 for each reaction. After the final reading, the mean value of all readings was calculated resulting in 24 different numbers ranging from 0 to 30 for each strain (the biochemical fingerprint). Similarities between the strains were calculated as correlation coefficients as described before [21] and clustered according to the unweighted pair group method with arithmetic averages (UPGMA) [22] yielding a dendrogram. An identity level of 0.975 was set based on reproducibility of the system as described before [23]. Strains showing similarities to each other higher than this value were regarded as identical and assigned to the same PhP types and those not identical to any other isolates were called single (Si) PhP types. Different PhP types were denoted by capital letters and those non-identical types belonging to the same cluster were denoted by the same capital letters followed by different numbers.

Phenotypic diversity of the bacterial populations within each serogroup was measured as Simpson's index of diversity (Di) [24]. Within each serogroup, biochemical homogeneity among the bacterial population was calculated as the mean of correlation coefficient between all isolates.

#### Comparison between typing methods

The agreement between the typing methods was measured as described earlier [25]. Briefly, the diversity of isolates was calculated by both typing methods. Members of each pair of isolates, which were either identical or different by both typing methods, yielded a score of +1 (which indicates a good agreement). Members of each pair, which were identical by the typing method showing the lowest diversity and different by the method showing the highest diversity, also yielded a score of +1 (good agreement). Finally, members of each pair, which were different by the typing method showing the lowest diversity and identical by the method showing the highest diversity, were given score of -1 (poor agreement). The sum of all scores was calculated and divided by the number of comparisons. By multiplying this value by 100, a value which is an estimate of the percentage agreement between two typing methods was obtained.

#### Statistical analysis

The  $\chi^2$  test was used for measuring the significance of differences.

#### Results

The most common serogroups among the EPEC strains studied were O127 and O114. Of the 23 EPEC strains

belonging to serogroup O127, 16 produced CDT. Results of both the PhP typing and the PCR typing indicated that most of these strains belonged to one common clonal group (i.e., PhP/PCR type C/b) (Fig. 1). The CDT-negative strains belonged to two different PhP/PCR types, one of which was closely related to CDT-positive strains (Fig. 1). To investigate the degree of relationship among these strains, phenotypic diversity and homogeneity were measured among them. A low diversity was found with both typing methods (i.e., 0.71 and 0.49 for PhP and PCR, respectively). There was also a high homogeneity (i.e., 0.78) among their biochemical phenotypes (Table 2).

Sixteen of the 19 strains of serogroup O114 belonged to the same PhP/PCR type (Fig. 2), thus showing a low diversity with both the PhP (0.30) and the PCR (0.30) typing methods, and a high homogeneity (0.87) with respect to their biochemical phenotypes (Table 2). A high agreement was obtained between the PhP and the PCR typing methods for strains of serogroups O127 (99%) and O114 (100%) (Table 2). According to both typing methods, EPEC strains belonging to less common serogroups ( $n = 21$ ) also belonged to many different clonal groups, although within certain serogroups there was a high similarity between the strains. For example, all five strains belonging to serogroup O125 had very similar PhP/PCR types. Similarly, three of the four strains belonging to serogroup O20 also had similar PhP/PCR types (Fig. 3). It was also found that some strains belonging to different serogroups (e.g., O20 and O126) had identical PhP or PCR types, or both (Fig. 3).

The O antigen non-typable (ONT) strains ( $n = 31$ ) were also highly diverse as shown by both methods, although some clusters of strains with similar PhP and PCR types were found (Fig. 4). For example, strains belonging to PhP types K and L (including L1) had similar PCR types (i.e., h and g, including g1 and g2, respectively) (Fig. 4). Again, isolates were found with different PhP types that had similar PCR types (e.g., PCR type e) (Fig. 4).

To find out whether ONT strains have originated from any of the known O serogroups, one isolate from each common and single PhP/PCR type shown in Figs. 1-3 was selected and compared with the ONT strains (Fig. 5). The results indicated that while only a few ONT strains were identical to strains of known serogroups (e.g., strains of PhP types N and L), most of the ONT strains were distantly related to each other and to the strains of known serogroups (Fig. 5). It was also found that strains with known serogroups were more related to each other than ONT strains.

#### Discussion

Infection due to EPEC continues to be a major cause of infantile diarrhoea in developing countries [8, 26].





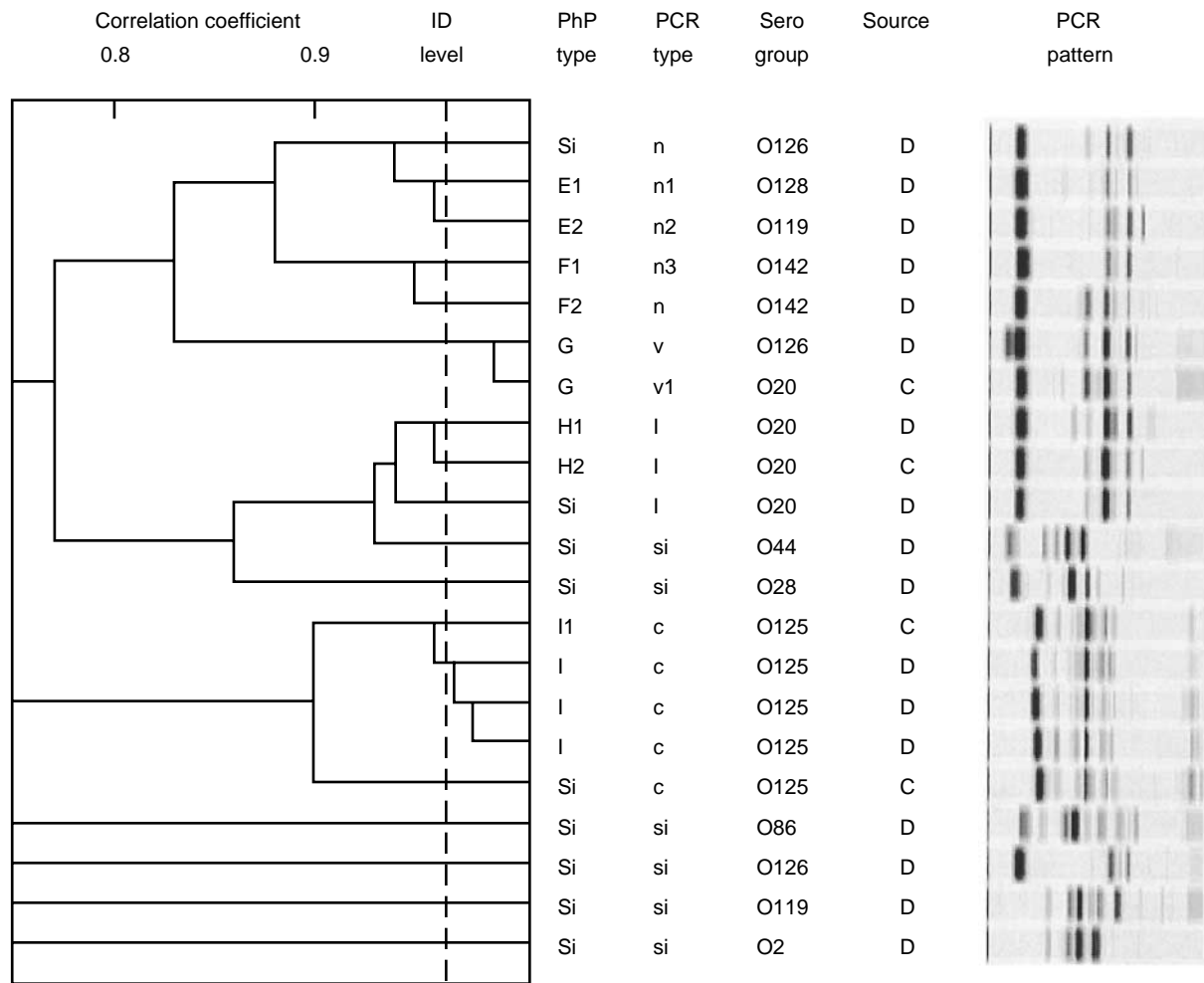


Fig. 3. UPGMA dendrogram derived from clustering of the PhP types and their comparison with PCR types of EPEC strains belonging to less prevalent serogroups isolated from case-control studies of diarrhoea in Bangladesh. D, diarrhoeal cases; C, healthy control; ID level, identity level; Si, PhP type found in only one isolate; si, PCR type found in only one isolate.

foreign to the *E. coli* chromosome. These regions are termed 'pathogenicity islands' which often harbour clustered virulence determinants, IS sequences and phage- or plasmid-like traits. These elements may act as vehicles for the horizontal transfer of virulence determinants giving rise to potential new pathogens [32]. The fact that CDT<sup>+</sup> strains in the present study had very similar PhP/PCR types to one of the two CDT<sup>-</sup> strains of the same serogroup suggests that CDT<sup>+</sup> strains might have evolved from CDT<sup>-</sup> strains by acquiring the *cdt* genes.

In determining the clonal origin of the EPEC isolates, a high agreement was found between the results obtained with all three typing methods used, although this was more pronounced with PhP and PCR typing than with either of these two methods and serogrouping. However, it should be noted that a good correlation between two typing methods does not necessarily mean that they are superior to other typing methods used in the same study, but simply implies that either method can be used when relations between the isolates are to be measured.

Twenty-one isolates belonging to other serogroups were also included in this study. Some of these serogroups belonged to class I EPEC and have been responsible for several outbreaks in different countries [33]. Results of the PhP typing indicated that they were highly diverse, sometimes even within a serogroup. For example, isolates belonging to O126 contained three distinct PhP types. On the other hand, with the PCR typing method, these isolates yielded the same genotypical pattern. Again, there was a higher agreement between the PhP and the PCR typing methods than between either of these methods and serogrouping.

In the present study a high percentage of EPEC isolates was also found which were not typable with traditional EPEC antisera and other available antisera. These ONT strains belonged to highly diverse PhP/PCR types [16]. Determination of virulence properties by DNA probes and cell culture techniques has improved the potential to identify true EPEC strains (i.e., strains with EAF and AE properties) in epidemiological studies [33–36]. All non-typable strains tested in the present study were shown to be true EPEC strains as determined by

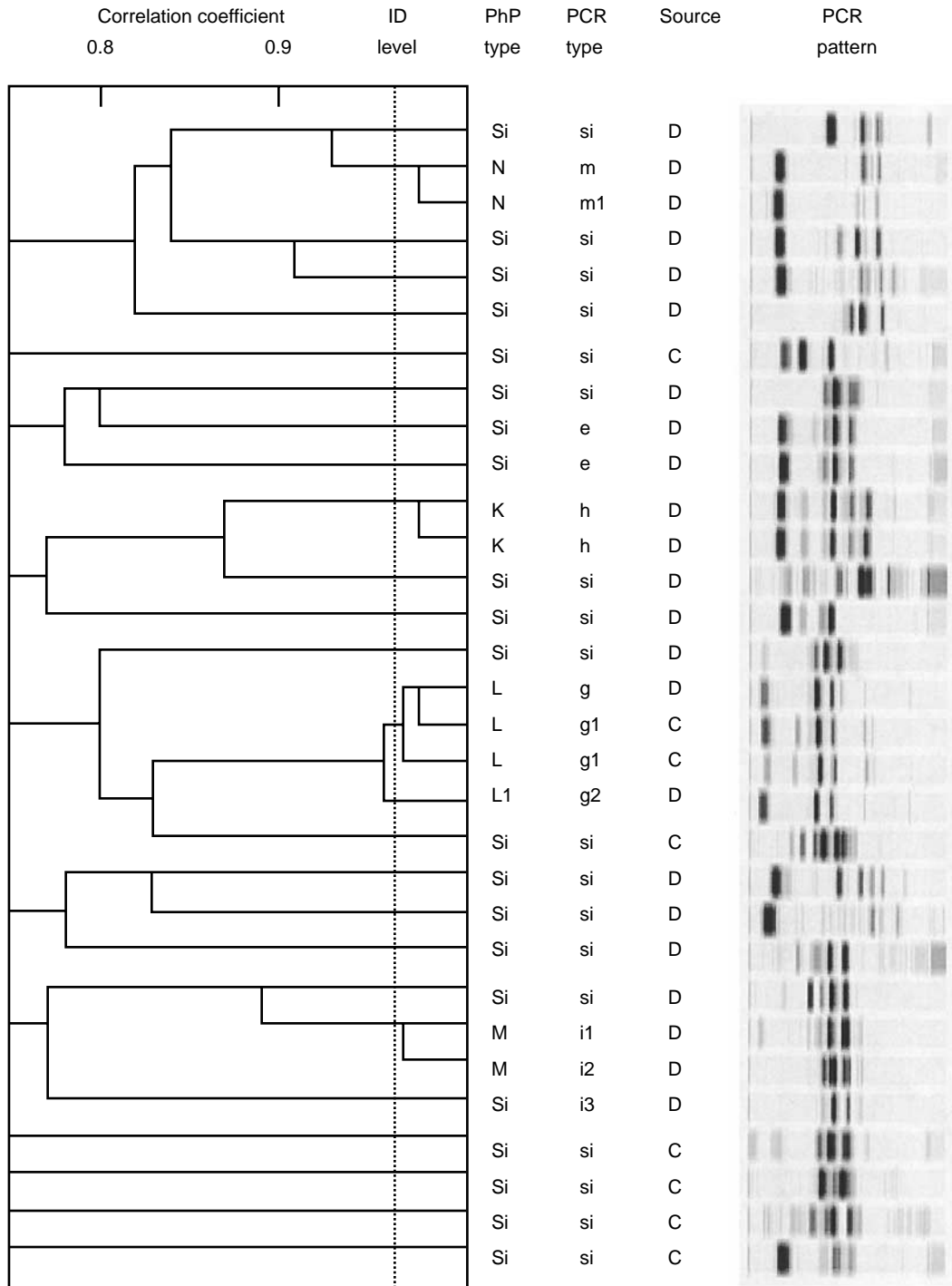


Fig. 4. UPGMA dendrogram derived from clustering of the PhP types and their comparison with PCR types of O antigen non-typable (ONT) EPEC strains, isolated from case-control studies of diarrhoea in Bangladesh. D, diarrhoeal case; C, healthy control; ID level, identity level; Si, PhP type found in only one isolate; si, PCR type found in only one isolate.

specific probes and cell culture assays. Non-typable EPEC strains with EAF and AE properties have occasionally been isolated from cases of diarrhoea [11, 37, 38], although their role as true pathogens in diarrhoeal disease is not known. The presence of such a high percentage of ONT EPEC strains in the present study is rather surprising. One possibility is that these strains had been true traditional EPEC serogroups and have gone through changes in their LPS content, or that they had been members of normal flora which have

acquired virulence properties, as exchange of genetic material among gut flora is common [39]. The fact that these strains were isolated with high frequency from both the patients and controls and also the finding that they were of highly diverse PhP/PCR types support this hypothesis. Moreover, most ONT strains did not show any phenotypic or genotypic similarity to any of the strains of the known serogroups, ruling out the possibility that they could represent newly emerged clones of EPEC in this country.

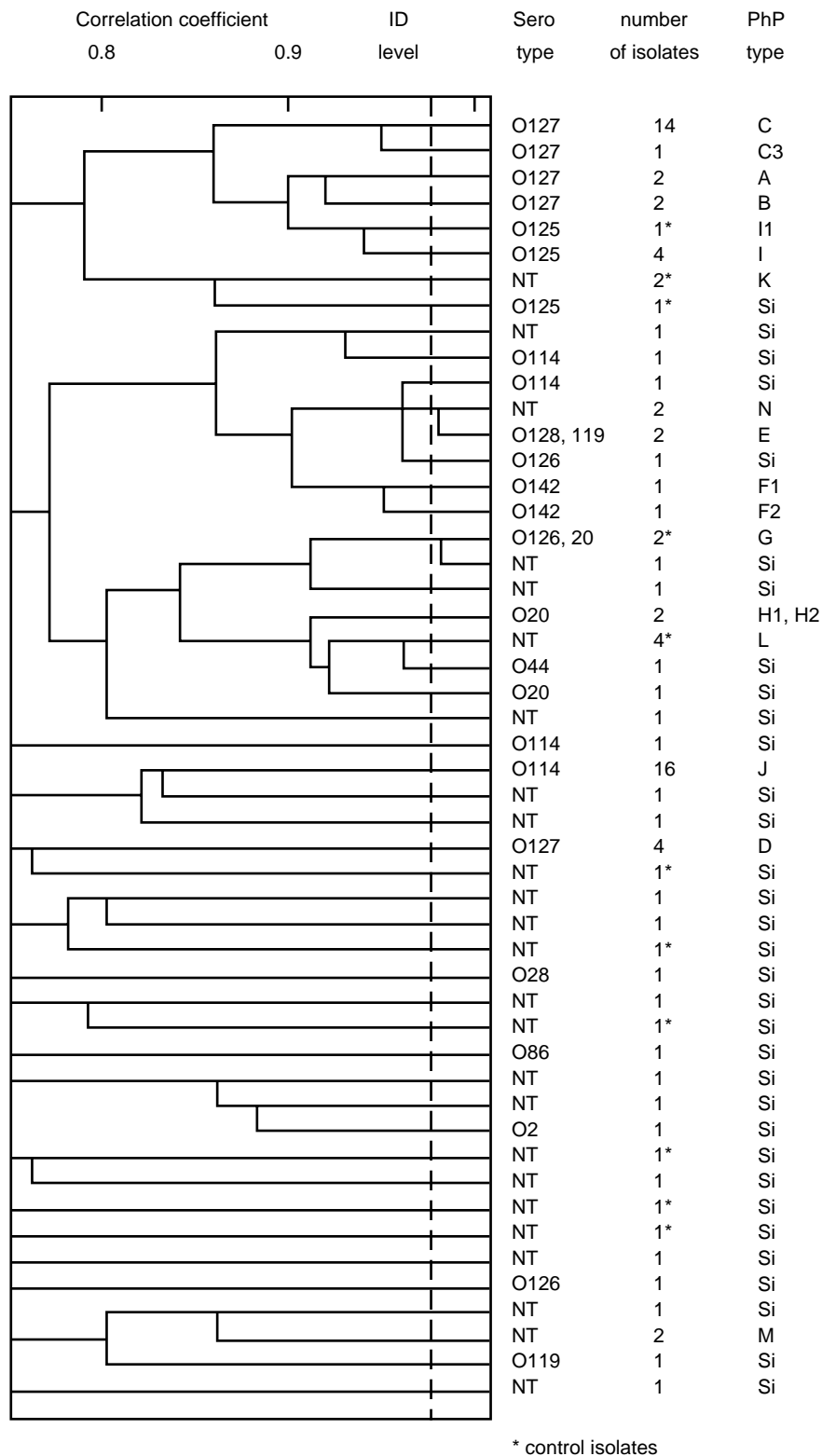


Fig. 5. UPGMA dendrogram derived from clustering of all PhP types identified in the present study, with the corresponding serotyping and PCR typing data added. One isolate representing each common and single PhP type (see Figs. 1–4) was included. ID level, identity level; ONT, O-antigen non-typable; Si, PhP type found in only one isolate; si, PCR type found in only one isolate.

In conclusion, two groups of EPEC strains, belonging to serogroups O114 and O127, were found to be responsible for the majority of the EPEC diarrhoea in Bangladesh. The prevalence of these serogroups in this country differs from that found in other parts of

the world. A high proportion of ONT EPEC strains was also found, which were of diverse clonal groups and probably have evolved from non-pathogenic clones of *E. coli* by acquiring the EPEC virulence determinants.

