

Molecular typing and virulence of enteroaggregative *Escherichia coli* strains isolated from children with and without diarrhoea in Rio de Janeiro city, Brazil

Adriana H. Regua-Mangia,¹ Tânia A. T. Gomes,² Mônica A. M. Vieira,² Kinue Irino³ and Lúcia M. Teixeira⁴

Correspondence

Adriana H. Regua-Mangia
regua@ensp.fiocruz.br

¹Departamento de Ciências Biológicas, Fundação Oswaldo Cruz (FIOCRUZ), Rio de Janeiro, Brazil

²Departamento de Microbiologia, Imunologia e Parasitologia, Universidade Federal de São Paulo (UNIFESP), São Paulo, Brazil

³Seção de Bacteriologia, Instituto Adolfo Lutz, São Paulo, Brazil

⁴Departamento de Microbiologia Médica, Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil

Enteroaggregative *Escherichia coli* (EAEC) strains have been implicated as emerging aetiological agents of diarrhoea worldwide. In the present study, 43 EAEC strains were serotyped and characterized according to random amplification of polymorphic DNA profiles, PFGE, multilocus enzyme electrophoresis (MLEE) and the presence of putative virulence genes (*hly*, *aero*, *kps*, *fim*, *aggA*, *aafA*, *aggR*, *astA*, *she*, *aap*, *shf* and *pet*). The EAEC strains consisted of a diversity of serotypes including eight O-non-typable and 35 O-typable strains arranged into 21 O:H combinations. Amplification of specific genes revealed that all strains carried at least two of the virulence sequences investigated. *fim*, *aggR* and *aap* were the most frequent genes in both groups studied. *hly*, *aero* and *aggA* sequences were more prevalent in the diarrhoeal group. *kps* occurred exclusively in strains isolated from symptomatic children and showed strong association with diarrhoeal disease. The molecular approaches used to investigate the relatedness among EAEC strains revealed a high degree of polymorphism, suggesting that these micro-organisms have a non-clonal origin. A closer relationship was observed among EAEC strains sharing O:H types. No significant clustering could be identified related to the virulence traits investigated; however, the *she* locus showed clonal distribution by MLEE typing. These results are in accordance with previous findings in revealing the conservation of particular EAEC factors, despite the high degree of diversity related to both genotypic and phenotypic markers.

Received 10 September 2008

Accepted 6 December 2008

INTRODUCTION

Enteroaggregative *Escherichia coli* (EAEC) strains have been implicated as emerging aetiological agents of diarrhoeal disease in both developing and developed countries, predominantly among cases that persist for longer than 14 days (Huang *et al.*, 2004; Regua-Mangia *et al.*, 2004a; Sarantuya *et al.*, 2004). They are characterized by the ability to adhere to cultured cell monolayers with an aggregative or 'stacked-brick' adhesion phenotype, and usually to the glass surface around the cells. The group comprises a diverse range of serotypes that possess a variety of putative virulence factors and express considerable heterogeneity in surface structures (Kahali *et al.*, 2004; Nataro & Kaper, 1998; Suzart *et al.*, 2001). The exact mechanisms by which

EAEC causes diarrhoea have not been completely defined. Most EAEC strains harbour a high-molecular-mass plasmid (designated pAA) from which a fragment is used as a detection probe (AA). pAA also encodes several putative virulence factors, including fimbrial adhesins designated AA fimbriae I, II and III (AAF/I–III) (Bernier *et al.*, 2002; Czeczulin *et al.*, 1997; Nataro *et al.*, 1993), the enteroaggregative heat-stable enterotoxin 1 (EAST-1) (Savarino *et al.*, 1991), a heat-labile enterotoxin designated Pet (Eslava *et al.*, 1998) and a dispersin surface coat protein (Aap, formerly known as AspU) (Sheikh *et al.*, 2002). The *E. coli* α -haemolysin (Hly) and the cytolethal distending toxin are other putative virulence factors detected in some EAEC strains (Albert *et al.*, 1996; Gomes *et al.*, 1995). In addition to pAA, some EAEC strains express putative factors encoded on the chromosome such as the 116 kDa secreted protein involved in intestinal colonization (Pic) (Henderson *et al.*, 1999) and a protein necessary for

Abbreviations: EAEC, enteroaggregative *Escherichia coli*; ET, electrophoretic type; MLEE, multilocus enzyme electrophoresis; ONT, O-non-typable; RAPD, random amplification of polymorphic DNA.

yersiniabactin biosynthesis (Irp2) (Schubert *et al.*, 1998). Although EAEC is considered an emerging category, not all strains are associated with diarrhoeal disease, and the category is heterogeneous regarding the presence of putative determinants (Nataro *et al.*, 1995; Suzart *et al.*, 1999, 2001). The role and prevalence of these factors have been found to vary by location, making the precise diagnosis of EAEC infection problematic (Huang *et al.*, 2004).

The increasing role of EAEC as an enteric pathogen in human infections, particularly among children in developing countries, has been the focus of much attention. Several investigations have been carried out in order to elucidate epidemiological aspects of EAEC infection, as well as to establish clearly the association of virulence markers with diarrhoeal disease (Nataro & Kaper, 1998; Regua-Mangia *et al.*, 2004a; Sarantuya *et al.*, 2004). Despite the heterogeneity commonly observed, the pathogenicity has been attributed to particular factors that are not uniformly distributed among EAEC strains isolated worldwide (Jenkins *et al.*, 2005; Nataro & Kaper, 1998; Okeke *et al.*, 2000; Sarantuya *et al.*, 2004; Spencer *et al.*, 1999; Uber *et al.*, 2006). Conventional methods have been used as simple tests of identification and epidemiological analysis of EAEC populations (Chart *et al.*, 1997; Sarantuya *et al.*, 2004; Suzart *et al.*, 1999, 2001). Molecular approaches, particularly DNA-based tests involving gene probes or PCR, have been applied successfully to elucidate pathogenic mechanisms, as well as to study the genetic diversity of this diarrhoeagenic *E. coli* category (Kahali *et al.*, 2004; Nataro & Kaper, 1998; Sarantuya *et al.*, 2004). Comparison of genetic markers among EAEC strains by molecular typing approaches has revealed the clonality of chromosomal sequences and a limited transfer of plasmid-borne genes (Boyd & Hartl, 1998; Kahali *et al.*, 2004). Such observations have been postulated to play a crucial role in the evolutionary process and the ability of this category to cause diarrhoea. DNA microarray analysis has shown that typical EAEC strains share a package of common virulence genes composed of a large number of conserved plasmid and chromosomal loci (Jenkins *et al.*, 2005).

In the present study, EAEC strains previously isolated from hospitalized children, with or without diarrhoea, in Rio de Janeiro city were characterized with regard to phenotypic and genotypic traits. The molecular relatedness among EAEC strains was evaluated by PFGE, random amplification of polymorphic DNA (RAPD) analysis and multilocus enzyme electrophoresis (MLEE).

METHODS

Bacterial strains. This study used 43 EAEC strains selected from a collection of 96 EAEC isolates originally obtained from 253 hospitalized children, less than 3 years of age, with and without diarrhoea, during a previous epidemiological investigation (Regua-Mangia *et al.*, 2004a). All strains were classified previously as EAEC by reactivity with the pCVD 432 EAEC probe in colony hybridization

assays and expression of the AA pattern in HEp-2 cells. A typical AA adherence pattern was used to classify EAEC strains expressing an aggregative or 'stacked-brick' pattern. An atypical pattern defined strains showing a non-typical AA phenotype mostly characterized by a chain-like adherence pattern or bacterial clusters adherent to epithelial cells or coverslips. Except for one strain that also carried the *eae* gene, these strains lacked DNA sequence similarity with other genetic markers described for diarrhoeagenic *E. coli* pathotypes. Bacterial cultures were kept at -70°C in trypticase soy broth (BBL; Becton Dickinson) with 20% glycerol.

Serotyping. Identification of somatic (O) and flagellar (H) antigens was performed by standard agglutination methods (Ewing, 1986), with the specific antisera O1–O181 and H1–H56 prepared in the Bacteriology Section of the Instituto Adolfo Lutz (São Paulo, Brazil). Non-motile strains were designated H⁻ and O-non-typable strains were defined as ONT.

MLEE. Bacterial cells were typed according to the variation of the following isoenzymes (Pupo *et al.*, 1997): aconitase, EC 4.2.1.3; adenylate kinase, EC 2.7.4.3; alcohol dehydrogenase, EC 1.1.1.1; glucose-6-phosphate dehydrogenase, EC 1.1.1.49; phosphogluconate dehydrogenase, EC 1.1.1.44; glucose phosphate isomerase, EC 5.3.1.9; isocitrate dehydrogenase, EC 1.1.1.42; malate dehydrogenase, EC 1.1.1.3.7; phenylalanyl-leucine peptidase, EC 3.4.11.1; and phosphoglucomutase, EC 2.7.5.1. Phenetic analysis based on isoenzyme data was carried out using the NTSYS-pc version 2.0 software package (F. James Rolf, Exeter Software) and the UPGMA algorithm. Isolates lacking detectable enzyme activity were designated as having a null state at the locus in question. Electrophoretic types (ETs) were defined based on the combinations of isoenzymic markers.

RAPD. RAPD analysis was performed as described previously (Regua-Mangia *et al.*, 2004b) using three arbitrary 10mer primers: 1254 (5'-CCGCAGCCAA-3'), 1253 (5'-GTTTCCGCC-3') and 1290 (5'-GTGGATGCGA-3'). Reaction products were analysed by electrophoresis on 1.5% agarose gels and stained with ethidium bromide. RAPD profiles were inspected visually and defined according to the presence or absence and intensity of polymorphic bands. A 1 kb DNA ladder was used as a molecular mass marker (Gibco-BRL). ETs were defined according to the presence or absence and intensity of polymorphic bands. The reproducibility of the RAPD amplifications was assessed using the selected primers with different DNA samples isolated independently from the same strain and amplified at different times.

Search for virulence DNA sequences related to EAEC. All strains were screened for the presence of *hly* (α -haemolysin), *aero* (aerobactin), *kps* (capsule), *fim* (type I fimbria), *aggA* (AAF/I fimbrial subunit), *aafA* (AAF/II fimbrial subunit), *aggR* (AAF/I and AAF/II transcriptional activator), *astA* (enteroaggregative heat-stable enterotoxin), *she* (secreted protein of *Shigella flexneri*), *aap* (plasmid-encoded surface coat protein), *shf* (cryptic ORF) and *pet* (enteroaggregative plasmid-encoded enterotoxin) genes by PCR (Okeke *et al.*, 2000). Bacterial strains were grown overnight at 37°C in trypticase soy agar (TSA) and PCR assays were conducted with bacterial DNA released from whole organisms by boiling. Assessment of the amplification efficiency and the precision of the reactions was carried out under optimized conditions and performed in duplicate on different assays. Each PCR analysis included a positive DNA control, a negative DNA control and a no-template control. Amplification products were analysed by electrophoresis on agarose gels and stained with ethidium bromide.

PFGE. The preparation of DNA for PFGE analysis was performed as described previously (Regua-Mangia *et al.*, 2008). Bacterial cells grown overnight on TSA plates were resuspended in PIV buffer [1 M

NaCl, 10 mM Tris/HCl (pH 7.6)] and embedded in 1% low-melting-point agarose (NuSieve GTG Agarose; FMC Bioproducts). Lysis of bacterial cells was carried out by placing plugs in a lysis solution containing 1 mg lysozyme ml⁻¹ followed by two steps of deproteinisation by washing with ESP solution [0.5 M EDTA (pH 8.0), 1% sodium lauroyl sarosine, 0.1% proteinase K ml⁻¹] for 18–24 h at 50 °C. The plugs were then washed seven times (four times for 1 h

each and three times for 2 h each) with TE buffer [10 mM Tris/HCl (pH 7.5), 0.1 mM EDTA] at 37 °C according to the manufacturer's instructions (Boehringer Mannheim). For restriction endonuclease digestion, DNA in the plugs was incubated with 20 U *Xba*I for 18–24 h at 37 °C. Electrophoresis was performed in 1% agarose gels (Molecular Biology Certified Agarose; Bio-Rad) in 0.5 × Tris/borate/EDTA buffer using the CHEF DR-III system (Bio-Rad). The

Table 1. General characteristics of the EAEC strains used in this study

Origin	Serotype	Adherence type	Putative virulence genes detected by PCR	Electrophoretic profiles				
				PFGE*	RAPD			MLEE
					1254	1253	1290	
Case	O105:H17	Atypical	<i>aero kps fim aggR aap shf</i>	2	2	2	2	1
	ONT:H18	Atypical	<i>kps fim aggA aggR aap</i>	3	3	3	3	2
	ONT:H18	Typical 6 h	<i>kps fim aggA aggR aap</i>	3	3	3	3	3
	O153:H2	Atypical	<i>aero kps fim aggA aggR she aap</i>	4	4	4	4	4
	O86:H ⁻	Atypical	<i>aero kps fim aggA aggR aap</i>	6	7	8	7	5
	O25:H17	Atypical	<i>aero kps fim</i>	7	8	9	8	6
	O86:H11	Atypical	<i>aero kps fim aggR aap</i>	8	5	5	5	7
	O86:H11	Atypical	<i>aero kps fim aggR aap</i>	8	5	6	5	7
	O86:H11	Atypical	<i>kps fim aggR aap</i>	9	9	10	9	7
	O15:H18	Atypical	<i>hly aero kps fim aggA aggR aap</i>	10	10	11	10	8
	O15:H18	Atypical	<i>hly aero fim aggA aggR she aap</i>	11	11	12	11	9
	O15:H18	Typical	<i>hly aero kps fim aggA aggR she aap</i>	12	10	13	12	10
	ONT:H10	Atypical	<i>aggR astA she aap</i>	13	12	14	13	11
	O73:H33	Atypical	<i>aero kps fim aggA aggR she aap</i>	D	13	15	13	12
	O44:H18	Typical 3 h	<i>aero kps fim aggA aggR she aap</i>	D	14	16	14	13
	O86:H11	Typical 3 h	<i>aero kps fim aggR aap</i>	16	19	21	19	14
	ONT:H32	Typical 6 h	<i>aero kps fim aap shf</i>	17	16	21	16	15
	ONT:H32	Typical 3 h	<i>aero kps fim aap</i>	17	16	18	16	15
	O113:H ⁻	Typical 3 h	<i>fim aggR astA she aap shf pet</i>	18	20	18	20	16
	O113:H ⁻	Typical 3 h	<i>fim aggR astA she aap shf pet</i>	18	20	22	20	16
	O3:H2	Typical 3 h	<i>hly aero kps fim aggA aggR astA she aap</i>	19	21	23	21	17
	O131:H ⁻	Atypical	<i>aero fim aggR she</i>	21	23	25	23	18
	O128:H10	Typical 3 h	<i>fim aggA aggR astA aap</i>	24	25	28	25	19
	O11:H10	Typical 3 h	<i>aero kps fim aggR astA aap shf pet</i>	25	26	29	25	20
	O73:H18	Atypical	<i>aero kps fim aggR she aap shf</i>	26	27	30	26	21
	O85:H10	Typical 3 h	<i>fim aggA aggR aap</i>	D	28	31	27	22
	O65:H ⁻	Typical 3 h	<i>fim aggR astA she aap</i>	27	28	32	27	23
	O86:H11	Typical 3 h	<i>aero kps fim aggR aap</i>	28	29	33	28	24
	O104:H17	Typical 3 h	<i>aero kps fim aggR astA she aap</i>	D	29	34	29	25
O3:H2	Typical 3 h	<i>hly aero kps fim aggA aggR aap shf</i>	29	30	35	30	26	
O3:H2	Typical 3 h	<i>hly aero kps fim aggA aggR aap shf</i>	29	30	35	30	26	
O8:H10	Typical 3 h	<i>fim aggA aggR aap</i>	30	31	36	31	27	
O65:H ⁻	Atypical	<i>fim astA she</i>	31	32	37	32	28	
O15:H18	Atypical	<i>hly aero kps fim aggA aggR she aap</i>	32	33	38	33	29	
Control	ONT:H10	Typical 3 h	<i>aero fim aggR aap shf pet</i>	1	1	1	1	30
	ONT:H ⁻	Typical 6 h	<i>aero fim aafA aggR astA she aap shf</i>	5	6	7	6	31
	O113:H12	Typical 3 h	<i>hly aero fim aggR she aap</i>	11	11	12	11	32
	ONT:H10	Typical 3 h	<i>aero fim aggR she aap shf</i>	14	15	17	15	33
	O20:H ⁻	Atypical	<i>fim aggR astA aap</i>	15	17	19	17	22
	O20:H ⁻	Typical 6 h	<i>hly aero fim aggR aap</i>	15	18	20	18	34
	O92:H33	Typical 3 h	<i>aero fim aggA aggR she aap</i>	20	22	24	22	17
	O113:H ⁻	Typical 3 h	<i>fim aggA aggR astA she aap shf pet</i>	23	24	27	25	16
	O113:H ⁻	Atypical	<i>fim aap</i>	22	24	26	24	28

*D, DNA degradation.

parameters were a pulse time ranging from 5 to 50 s for 23 h at 11 °C, in a voltage gradient of 6 V cm⁻¹. A 50–1000 kb ladder (Sigma) was used as a molecular mass marker. PFGE profiles were compared by visual and automated analysis using the Molecular Analyst Fingerprint Plus software (Bio-Rad), version 1.12. Cluster analysis was done using the UPGMA method of the Image Analysis System applying a maximum tolerance of 1.2 %, and the percentage similarity was estimated using the Dice coefficient.

Statistical analysis. The association between putative virulence genes and EAEC isolates obtained from children in the patient group and in the control group was analysed and compared by χ^2 and Fisher's exact tests using the Epi Info program, version 3.5.1. *P* values less than 0.05 were considered significant.

RESULTS

The general characteristics of the EAEC strains included in the present study are shown in Table 1. A high diversity of serotypes was identified (Table 1; Table 2). The 43 EAEC strains consisted of eight ONT and 35 O-typable strains. The ONT EAEC included one non-motile (H⁻) strain and the flagellar types H18 (25 %), H10 (38 %) and H32 (25 %). The O-typable strains comprised 18 O-groups, eight H-types and five non-motile strains arranged into 21 serotypes. The most frequent serotypes included in this study were: O86:H11 (11.6 %); O15:H18 and O113:H⁻ (9.3 % each); ONT:H10 and O3:H2 (6.9 % each); and O20:H⁻, O65:H⁻, ONT:H18 and ONT:H32 (4.6 % each). EAEC strains belonging to the same serotype often had distinct AA pattern. Serotypes O3:H2, O15:H18,

O65:H⁻, O86:H11, ONT:H18 and ONT:H32 were found only in the patient group. Serotypes ONT:H10 and O113:H⁻ were shared among EAEC strains isolated from diarrhoeic and control children exhibiting typical and atypical adherence phenotypes. Serotypes O20:H⁻, O92:H3, O113:H12 and ONT:H⁻ were identified only among control children and were related to both typical and atypical AA patterns. EAEC serotypes detected only among diarrhoeic children were found to exhibit the typical AA pattern.

Results from amplification reactions of chromosomal and plasmid-encoded genes showed that all strains carried at least two of the virulence determinants searched for (Table 1). The prevalence of chromosomal and pAA-encoded factors varied among case and control children. The frequency of distinct virulence genes was slightly higher among strains recovered from children with diarrhoea than from asymptomatic controls and was found in strains expressing both typical and atypical adherence phenotypes (Table 1; Table 2). *hly*, *aero* and *aggA* sequences were more prevalent in the diarrhoeal group, detected in 20.6, 67.6 and 47.0 % of the strains isolated from this group, respectively (Table 2). The *kps* gene occurred exclusively among symptomatic children (70.6 %). The *aafA* sequence, which encodes AAF/II pilin, was detected in only one strain from a healthy child. Some specific gene combinations occurred more frequently among children with diarrhoea: *hly-aero-fim-aggR-aap*; *fim-aggA-aggR-aap*; *fim-aggR-aap-shf-pet*; *aero-fim*; *kps-fim*; and *fim-aap-shf*.

Table 2. Overall distribution of putative virulence genes in EAEC isolates from serotypes associated with diarrhoeic and non-diarrhoeic children

Of the total 43 isolates, 34 were from diarrhoeal disease and nine were obtained from control subjects. The numbers in the columns represent the number of isolates. The first number in parentheses indicates the percentage of the total category positive for the respective characteristic. The second number for the third column indicates the *P* value for χ^2 or Fisher's test comparing isolates from diarrhoeic children with those in the non-diarrhoeic group. Serotypes ONT:H10 and O113:H⁻ were found in both groups.

Putative virulence gene detected by PCR	Distribution		
	Total (%)	Patient group (%; <i>P</i> value)	Control group (%)
<i>hly</i>	9 (20.9)	7 (20.6; 0.7890)	2 (22.2)
<i>aero</i>	29 (67.4)	23 (67.6; 0.623)	6 (66.7)
<i>kps</i>	24 (55.8)	24 (70.6; 0.0001)	0 (0)
<i>fim</i>	42 (97.7)	33 (97; 0.790)	9 (100)
<i>aggA</i>	18 (41.9)	16 (47; 0.168)	2 (22.2)
<i>aafA</i>	1 (2.3)	0 (0; 0.209)	1 (11.1)
<i>aggR</i>	38 (88.4)	30 (88.2; 0.722)	8 (88.9)
<i>astA</i>	12 (27.9)	9 (26.5; 0.488)	3 (33.3)
<i>she</i>	20 (46.5)	15 (44.1; 0.405)	5 (55.5)
<i>aap</i>	40 (93.0)	31 (91.2; 0.484)	9 (100)
<i>shf</i>	12 (27.9)	8 (23.5; 0.201)	4 (44.4)
<i>pet</i>	2 (22.2)	3 (8.8; 0.405)	2 (22.2)
No. of isolates	43	34	9
No. of serotypes	27	23	6

The *aafA* and *aggA* genes were mutually exclusive and, in general, occurred along with the *aggR* sequence. EAEC isolates within a serotype were not homogeneous in terms of virulence patterns.

The molecular relatedness among EAEC strains was investigated by PFGE, RAPD and MLEE methodologies. PFGE analysis using *Xba*I endonuclease generated 32 distinct electrophoretic profiles composed of 12–17 polymorphic bands ranging from 50 to 580 kb (Fig. 1). PFGE typing revealed a genetically diverse population with a Dice similarity index ranging from 45 to 93%. Closer relatedness was observed especially among EAEC strains sharing O:H types. PFGE clustering showed defined clonal groups, but no clear correlation was observed with the virulence markers investigated. Although generating stable, reproducible and discriminatory banding profiles, PFGE analysis could not be performed in strains belonging to specific serotypes such as O44:H18, O73:H33, O85:H10 and O104:H17 due to DNA degradation. These isolates were, in turn, typed and discriminated by RAPD analysis. RAPD assays generated electrophoretic profiles composed of 5–21 bands ranging from approximately 300 to 3000 bp (Fig. 2). Reactions performed with primers 1254 and 1253 (70 mol% G+C) and 1290 (60 mol% G+C) resulted in 33, 38 and 33 different RAPD profiles, respectively (Table 1). The total number of polymorphic bands was 21, 16 and 15 for the 1254, 1253 and 1290 primers, respectively. There was a direct correlation between higher G+C content, the ability of the primer to detect polymorphism, and the reproducibility and stability of the reactions. MLEE data showed that all isoenzymes investigated were polymorphic, with the number of electromorphs ranging from two for isocitrate dehydrogenase, glucose phosphate isomerase and aconitase to five for phosphogluconate dehydrogenase and adenylate kinase. The most polymorphic isoenzymes were phosphogluconate dehydrogenase, adenylate kinase, phosphoglucomutase and malate dehydrogenase. Comparisons of the zymovars revealed 34 distinct zymovars or MLEE types (Table 1). Four ETs were represented by single serotypes and three ETs included multiple serotypes. The dendrogram derived from MLEE typing defined two main clusters (Fig. 3). Cluster A was subdivided into A1 and A2 subgroups composed of strains from diverse serotypes and virulence traits ranging from 80 to 100% similarity. EAEC strains isolated from non-diarrhoeic children were grouped in subgroup A2. Cluster B represented the major cluster, with 24 EAEC strains arranged into three subclusters (B1, B2 and B3) in separate branches or small clonal groups as observed with cluster A. The two clonal complexes were characterized by serotype-specific clonal cluster distribution. No significant clustering was related to the virulence traits investigated, which were dispersed widely throughout the MLEE tree. Additionally, it should be noted that the chromosomal *she* locus was not grouped in clusters B1 or B2.

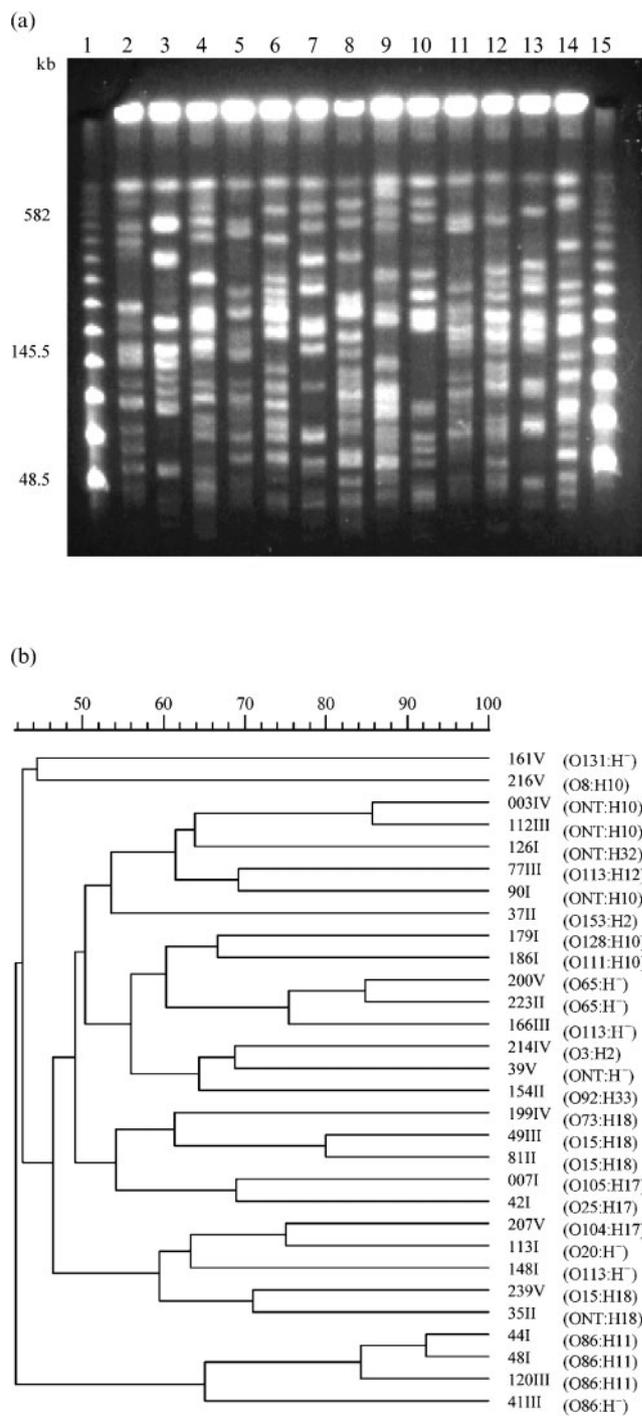


Fig. 1. (a) Representative PFGE profiles of *Xba*I-digested DNA of EAEC. Lanes: 1 and 15, molecular mass markers (lambda DNA concatemers ranging from 48.5 to 1018 kb); 2, strain 003IV; 3, strain 116II; 4, strain 214IV; 5, strain 223III; 6, strain 199IV; 7, strain 216II; 8, strain 81I; 9, strain 39V; 10, strain 48III; 11, strain 200V; 12, strain 152I; 13, strain 37III; 14, strain 207IV. (b) Dendrogram of PFGE profiles generated by the Dice co-efficient and clustering by UPGMA.

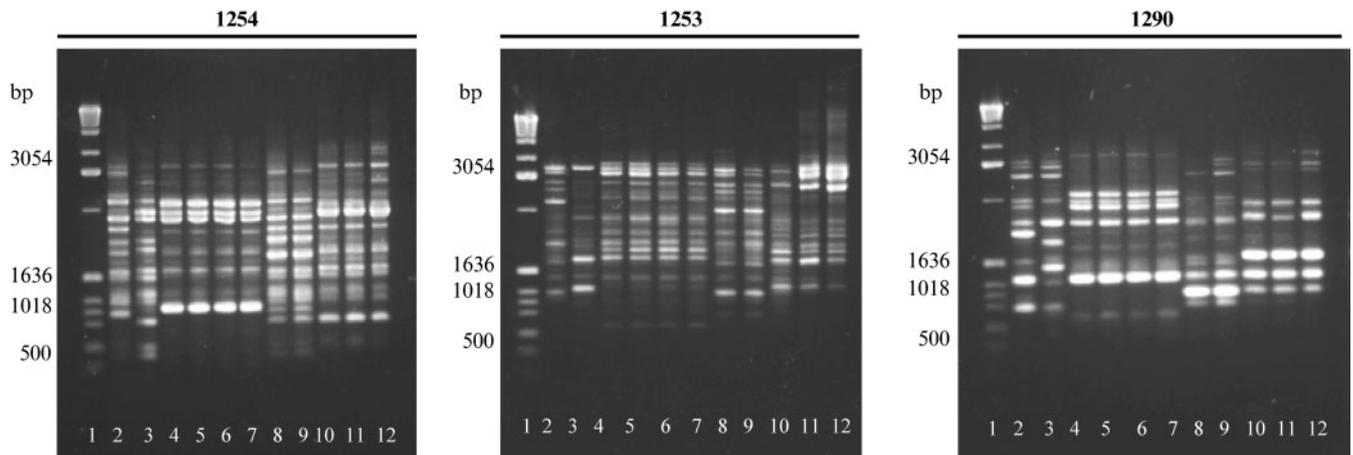


Fig. 2. RAPD profiles of EAEC strains obtained using three different primers (1254, 1253 and 1290). Lanes: 1, 1 kb DNA ladder; 2, strain 003IV; 3, strain 007I; 4, strain 034IV; 5, strain 035II; 6, strain 035III; 7, strain 035V; 8, strain 37II; 9, strain 37III; 10, strain 44I; 11, strain 44III; 12, strain 44IV.

DISCUSSION

EAEC is a human diarrhoeal pathogen of emerging importance, yet its virulence factors and their phylogenetics have not been completely characterized. In accordance with previous observations (Nataro & Kaper, 1998; Okeke *et al.*, 2000; Regua-Mangia *et al.*, 2004a; Sarantuya *et al.*, 2004), the present investigation revealed a significant EAEC strain-to-strain diversity comprising a heterogeneous set of pathogens that share particular genotypic and phenotypic markers.

A diversity of serotypes was detected in the studied population. Among the 23 serotypes that were recovered from children with diarrhoea, serotype O86:H11 was the most prevalent (11.6%). Some investigations have shown that strains belonging to serogroup O44 and some serotypes of the traditional enteropathogenic *E. coli* and enterotoxigenic *E. coli* serogroups such as O55, O86, O111, O128 and O126 are occasionally found among EAEC strains (Elias *et al.*, 2002; Nataro & Kaper, 1998; Suzart *et al.*, 1999), whilst other serotypes belonging to the O86, O126 and O125 serogroups are predominantly EAEC strains (Sarantuya *et al.*, 2004; Spencer *et al.*, 1999; Uber *et al.*, 2006; Valle *et al.*, 1997). In the present study, this relationship was also observed. Additionally, the serotypes more frequently detected have been found as the most common EAEC serotypes isolated in different geographical areas (Kahali *et al.*, 2004; Uber *et al.*, 2006). As with the findings of previous studies, we observed a high frequency of EAEC strains that expressed untypable O or H antigens or were non-motile, as well as the prevalence of H18 and H2 antigens, particularly associated with diarrhoeal disease (Kahali *et al.*, 2004; Nataro & Kaper, 1998; Uber *et al.*, 2006).

The terms typical and atypical EAEC have been suggested to refer to strains harbouring or lacking the *aggR* regulon, respectively (Huang *et al.*, 2004). Plasmid and chro-

mosomal EAEC virulence genes have been found to be under the control of *aggR*, which led to the recognition of this marker for truly virulent EAEC strains. In our study, the *aggR* gene was detected in strains from both diarrhoeic (88%) and non-diarrhoeic (78%) children, showing typical and atypical AA adherence patterns, differing from some previous epidemiological studies (Sarantuya *et al.*, 2004).

The investigation of virulence determinants revealed that the *hly*, *aero*, *aggA* and *kps* genes were more frequent in EAEC isolated from diarrhoeic than from non-diarrhoeic children; however, χ^2 and Fisher's tests showed that only the *kps* gene was significantly associated with diarrhoea ($P < 0.001$) (Table 2). *aafA* was detected in only one isolate from a non-diarrhoeic child. Despite presenting a high prevalence among isolates in the diarrhoeal group (47%), the *aggA* gene lacked significance with disease ($P > 0.05$). These findings agree with earlier studies in the USA (Sheikh *et al.*, 2002), but differed somewhat from others carried out in Brazil (Gioppo *et al.*, 2000), suggesting that regional aspects may be involved in the prevalence of EAEC fimbrial adhesins. Therefore, these results do not support the suggestion that AAF/II is a useful reference marker for detection of EAEC pathogenic strains worldwide (Okeke *et al.*, 2000). The *aggR* gene was present in several cases in the absence of the fimbrial subunit genes, raising the possibility of uncharacterized AAFs or other fimbriae not investigated here. Previously, the *aap* gene was found to be a more sensitive marker to identify this category than the EAEC probe sequence (Spencer *et al.*, 1999). In the present study, *aap* was detected in almost all EAEC probe-positive strains; however, this marker would not be suitable for the identification of virulent EAEC subpopulations when its occurrence among EAEC strains exhibiting a non-typical AA pattern and the fact that it is widely distributed among strains isolated from non-diarrhoeic children are consid-

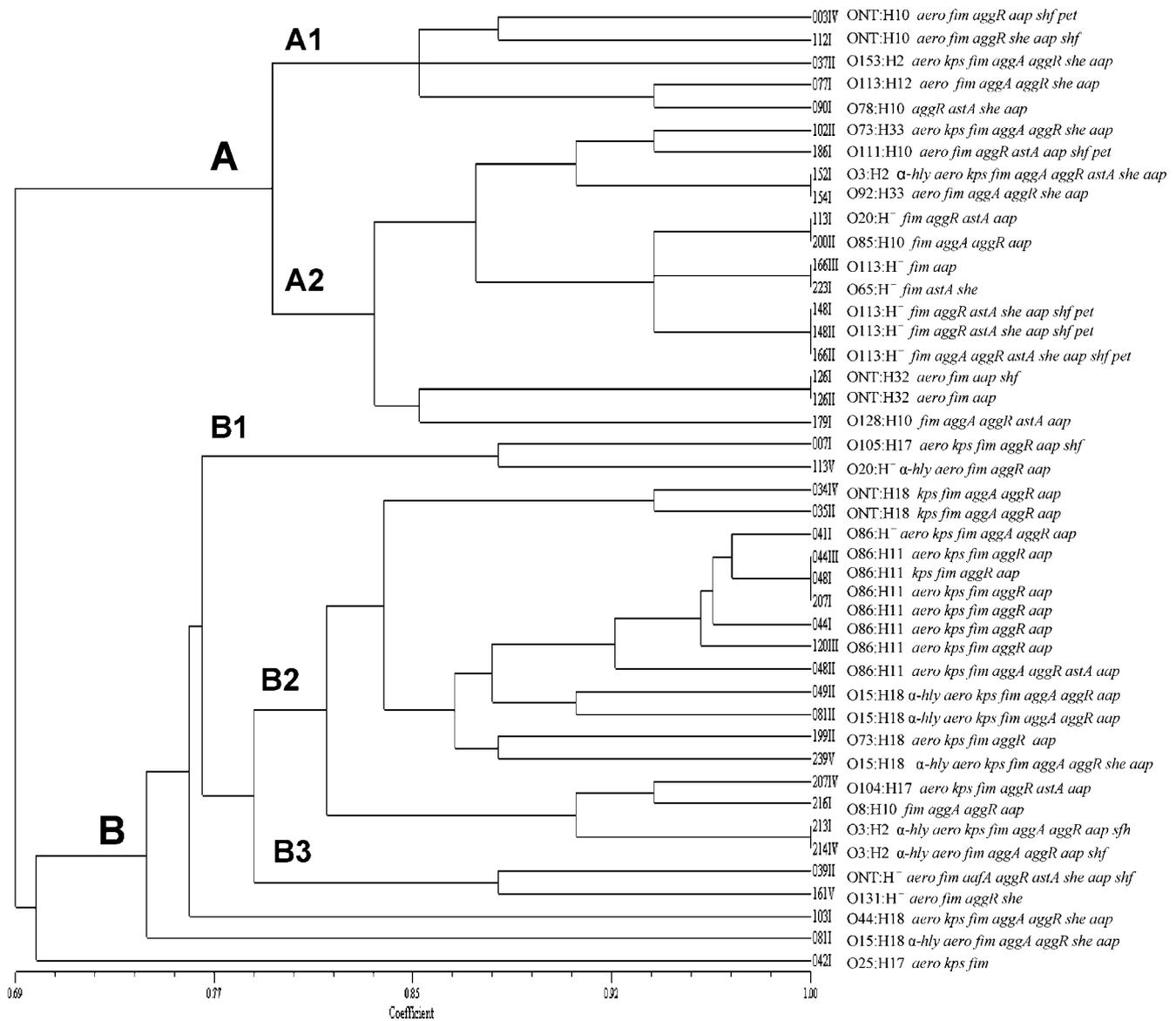


Fig. 3. Isoenzymic relationships and virulence among EAEC strains.

ered. Haemolysin production has been closely related to cytodetachment activity (Gomes *et al.*, 1995), and the locus encoding α -haemolysin has been detected largely among CD/EAEC probe-positive strains (Suzart *et al.*, 2001). In contrast to these observations, we detected the *hly* gene in only a minority of strains from both diarrhoeal and control children exhibiting a typical AA pattern. The lack of statistical significance of the *pet*, *astA* and *shf* genes with diarrhoeal disease ($P > 0.05$) and the low correlation with relevant characteristics do not suggest a potential role for the corresponding proteins in EAEC enteropathogenicity. The *kps* gene was found in 70.6% of the diarrhoeal group and was strongly associated with disease, but was absent in samples isolated from controls, suggesting that such a target may be a possible marker for epidemiological and

pathogenic purposes in our community. PCR methodology is a recognized powerful tool for the characterization of bacterial populations, providing reproducibility, sensitivity and specificity using well-standardized protocols. The inability of specific primers to amplify putative virulence genes from EAEC may be due to strains presenting mutations at the primer-binding sites or to genes encoding distinct sequences.

The molecular approaches used to investigate the overall chromosomal relatedness among EAEC strains were strongly correlated and showed a high genetic diversity, suggesting that these micro-organisms are panmictic. Such genotypic diversity may be explained by differences in the time of isolation, geographical origin and O:H types

(Arbeit *et al.*, 1990). Although EAEC clustering by comparing the three molecular typing methods used here was in broad agreement, the correlation was not complete. In contrast to isoenzymic characterization in which analysis conserves metabolic enzymes, PFGE and RAPD approaches provide information that is often related to specific loci where conservation, nature and the mode of evolution are unknown (Feil *et al.*, 2001). Thus, considering that these methodologies rely on distinct target analysis, it can be supposed that these micro-organisms constitute epidemiologically independent strains and may represent distinct evolutionary lineages. Despite the genetic variability, clustering analysis revealed that the degree of diversity was less among strains sharing O:H types. The recognition of serotype as a marker for clonality within *E. coli* populations has been supported since early studies on the aetiology and epidemiology of diarrhoeal disease. In our study, MLEE analysis showed that some strains belonging to distinct serotypes shared isoenzymic ETs. Such findings lead us to believe that the investigation of additional enzymic systems is required in order to infer the precise relationships among such EAEC strains.

Dendrogram analysis based on virulence profiles showed a widely dispersed distribution, even among EAEC strains that were distantly related, suggesting extensive lateral gene transference (Czeczulin *et al.*, 1999). Considering that most of the virulence markers investigated are plasmid-borne genes, which, in turn, can be lost or acquired horizontally, this favours the emergence of a variety of genotypic profiles composed of several combinations of virulence traits (Czeczulin *et al.*, 1999). Previous observations have suggested that the origin of EAEC strains may be derived from unrelated or commensal *E. coli* through independent acquisition of virulence factors (Czeczulin *et al.*, 1999; Gomes *et al.*, 1995; Pupo *et al.*, 1997).

Many of the results analysed were in accordance with previous findings regarding the conservation of particular EAEC factors, despite the high degree of diversity related to both genotypic and phenotypic markers. The role of these factors is fairly well defined, but requires additional investigations. The potential usefulness of the *kps* gene as a genetic marker for the tracking of potential virulent EAEC strains circulating in our community was identified.

ACKNOWLEDGEMENTS

This work was supported by a grant from Fundação Oswaldo Cruz (Programa de Apoio à Pesquisa Estratégica em Saúde, PAPES 2 – 250.250.339) and from Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ E-26/170.644/05, APQ1), Brazil.

REFERENCES

Albert, M. J., Faruque, S. M., Faruque, A. S. G., Betelheim, K. A., Neogi, P. K. B., Bhuiyan, N. A. & Kaper, J. B. (1996). Controlled study of cytotoxic distending toxin-producing *Escherichia coli* infections in Bangladeshi children. *J Clin Microbiol* **34**, 717–719.

- Arbeit, R. D., Arthur, M., Dunn, R., Kim, C., Selander, R. K. & Goldstein, R. (1990). Resolution of recent evolutionary divergence among *Escherichia coli* from related lineages: the application of pulsed field electrophoresis to molecular epidemiology. *J Infect Dis* **161**, 230–235.
- Bernier, C., Gounon, P. & Le Bouguenec, C. (2002). Identification of an aggregative adhesion fimbria (AAF) type III-encoding operon in enteroaggregative *Escherichia coli* as a sensitive probe for detecting the AAF-encoding operon family. *Infect Immun* **70**, 4302–4311.
- Boyd, E. F. & Hartl, D. L. (1998). Chromosomal regions specific to pathogenic isolates of *Escherichia coli* have a phylogenetically clustered distribution. *J Bacteriol* **180**, 1159–1165.
- Chart, H., Spencer, J., Smith, H. R. & Rowe, B. (1997). Identification of enteroaggregative *Escherichia coli* based on surface properties. *J Appl Microbiol* **83**, 712–717.
- Czeczulin, J. R., Balepur, S., Hicks, S., Philips, A., Hall, R., Kothary, M. H., Navarro-García, F. & Nataro, J. P. (1997). Aggregative adherence fimbria II, a second fimbrial antigen mediating aggregative adherence in enteroaggregative *Escherichia coli*. *Infect Immun* **65**, 4135–4145.
- Czeczulin, J. R., Whittam, T. S., Henderson, I. R., Navarro-García, F. & Nataro, J. P. (1999). Phylogenetic analysis of enteroaggregative and diffusely adherent *Escherichia coli*. *Infect Immun* **67**, 2692–2699.
- Elias, W. P., Barros, S. F., Moreira, C. G., Trabulsi, R. & Gomes, T. A. T. (2002). Enteroaggregative *Escherichia coli* strains among classical enteropathogenic *Escherichia coli* O serogroups. *J Clin Microbiol* **40**, 3540–3541.
- Eslava, C., Navarro-García, F., Czeczulin, J. R., Henderson, I. R., Cravioto, A. & Nataro, J. P. (1998). Pet, an autotransporter enterotoxin from enteroaggregative *Escherichia coli*. *Infect Immun* **66**, 3155–3163.
- Ewing, W. H. (1986). The genus *Escherichia*. In *Edwards & Ewing's Identification of Enterobacteriaceae*, 4th edn, pp. 93–134. Edited by P. R. Edwards & W. H. Ewing. New York: Elsevier.
- Feil, E. J., Holmes, E. C., Bessen, D. E., Chan, M. S., Day, N. P. J., Enright, M. C., Goldstein, R., Hood, D. W., Kalia, A. & other authors (2001). Recombination within natural populations of pathogenic bacteria: short-term empirical estimates and long-term phylogenetic consequences. *Proc Natl Acad Sci U S A* **98**, 182–187.
- Gioppo, N. M. R., Elias, W. P., Vidotto, M. C., Linhares, R. E. L., Saridakis, H. O., Gomes, T. A. T., Trabulsi, L. R. & Pelayo, J. S. (2000). Prevalence of HEp-2 cell-adherent *Escherichia coli* and characterization of enteroaggregative *E. coli* and chain-like adherent *E. coli* isolated from children with and without diarrhoea, in Londrina, Brazil. *FEMS Microbiol Lett* **190**, 293–298.
- Gomes, T. A. T., Abe, C. M. & Marques, L. R. M. (1995). Detection of HeLa cell-detaching activity and alpha-hemolysin production in enteroaggregative *Escherichia coli* strains isolated from feces of Brazilian children. *J Clin Microbiol* **33**, 3364.
- Henderson, I. R., Czeczulin, J., Eslava, C., Noriega, F. & Nataro, J. P. (1999). Characterization of Pic, a secreted protease of *Shigella flexneri* and enteroaggregative *Escherichia coli*. *Infect Immun* **67**, 5587–5596.
- Huang, D. B., Okhuysen, P. C., Jiang, Z. D. & DuPont, H. L. (2004). Enteroaggregative *Escherichia coli*: an emerging enteric pathogen. *Am J Gastroenterol* **99**, 383–389.
- Jenkins, C., Ijperen, C., Dudley, E. G., Chart, H., Willshaw, G. A., Cheasty, T., Smith, H. R. & Nataro, J. P. (2005). Use of a microarray to assess the distribution of plasmid and chromosomal virulence genes in strains of enteroaggregative *Escherichia coli*. *FEMS Microbiol Lett* **253**, 119–124.
- Kahali, S., Sarkar, B., Rajendran, K., Khanam, J., Yamasaki, S., Nandy, R. K., Bhattacharya, S. K. & Ramamurthy, T. (2004).

Virulence characteristics and molecular epidemiology of enteroaggregative *Escherichia coli* isolates from hospitalized diarrheal patients in Kolkata, India. *J Clin Microbiol* **42**, 4111–4120.

Nataro, J. P. & Kaper, J. B. (1998). Diarrheagenic *Escherichia coli*. *Clin Microbiol Rev* **11**, 142–201.

Nataro, J. P., Yikang, D., Girón, J. A., Savarino, S. J., Kothary, M. H. & Hall, R. (1993). Aggregative adherence fimbria I expression in enteroaggregative *Escherichia coli* requires two unlinked plasmid regions. *Infect Immun* **61**, 1126–1131.

Nataro, J. P., Yikang, D., Cookson, S., Cravioto, A., Savarino, S. J., Guers, L. D., Levine, M. M. & Tacket, C. O. (1995). Heterogeneity of enteroaggregative *Escherichia coli* virulence demonstrated in volunteers. *J Infect Dis* **171**, 465–468.

Okeke, I. N., Lamikanra, A., Czezulín, J., Dubovsky, F., Kaper, J. B. & Nataro, J. P. (2000). Heterogeneous virulence of enteroaggregative *Escherichia coli* strains isolated from children in Southwest Nigeria. *J Infect Dis* **181**, 252–260.

Pupo, G. M., Karaolis, D. K. R., Lan, R. & Reeves, P. R. (1997). Evolutionary relationship among pathogenic and nonpathogenic *Escherichia coli* strains inferred from multilocus enzyme electrophoresis and *mdh* sequence studies. *Infect Immun* **65**, 2685–2692.

Regua-Mangia, A. H., Gomes, T. A. T., Vieira, M. A., Andrade, J. R. C., Irino, K. & Teixeira, L. M. (2004a). Frequency and characteristics of diarrheagenic *Escherichia coli* strains isolated from children with and without diarrhea in Rio de Janeiro, Brazil. *J Infect* **48**, 161–167.

Regua-Mangia, A. H., Guth, B. E. C., Andrade, J. R. C., Irino, K., Pacheco, A. B., Ferreira, L. C., Zahner, V. & Teixeira, L. M. (2004b). Genotypic and phenotypic characterization of enterotoxigenic *Escherichia coli* (EPEC) strains isolated in Rio de Janeiro city, Brazil. *FEMS Immunol Med Microbiol* **40**, 155–162.

Regua-Mangia, A. H., Andrade, J. R., Gonzalez, A. G., Zahner, V., Cerqueira, A. M. & Teixeira, L. M. (2008). Genetic relatedness of a non-motile variant O157 enteropathogenic *Escherichia coli* (EPEC) strain and *E. coli* strains belonging to pathogenic related groups. *Microbiol Res* **163**, 225–233.

Sarantuya, J., Nishi, J., Wakimoto, N., Erdene, S., Nataro, J. P., Sheikh, J., Iwashita, M., Manago, K., Tokuda, K. & other authors (2004). Typical enteroaggregative *Escherichia coli* is the most prevalent pathotype among *E. coli* strains causing diarrhea in Mongolian children. *J Clin Microbiol* **42**, 133–139.

Savarino, S. J., Fasano, A., Robertson, D. C. & Levine, M. M. (1991). Enteroaggregative *Escherichia coli* elaborate a heat-stable enterotoxin demonstrable in an in vitro rabbit intestinal model. *J Clin Invest* **87**, 1450–1455.

Schubert, S., Rakin, A., Karch, H., Carniel, E. & Heeseman, J. (1998). Prevalence of the “high-pathogenicity island” of *Yersinia* species among *Escherichia coli* strains that are pathogenic for humans. *Infect Immun* **66**, 480–485.

Sheikh, J., Czezulín, J. R., Harrington, S., Hicks, S., Henderson, I. R., Le Bouguenec, C., Gounon, P., Philips, A. & Nataro, J. P. (2002). A novel dispersin protein in enteroaggregative *Escherichia coli*. *J Clin Invest* **110**, 1329–1337.

Spencer, J., Smith, H. R. & Chart, H. (1999). Characterization of enteroaggregative *Escherichia coli* isolated from outbreaks of diarrhoeal disease in England. *Epidemiol Infect* **123**, 413–421.

Suzart, S., Gomes, T. A. T. & Guth, B. E. C. (1999). Characterization of serotypes and outer membrane protein profiles in enteroaggregative *Escherichia coli* strains. *Microbiol Immunol* **43**, 201–205.

Suzart, S., Guth, B. E. C., Pedroso, M. Z., Okafor, U. M. & Gomes, T. A. T. (2001). Diversity of surface structures and virulence genetic markers among enteroaggregative *Escherichia coli* (EAEC) strains with and without the EAEC DNA probe sequence. *FEMS Microbiol Lett* **201**, 163–168.

Uber, A. P., Trabulsi, L. R., Irino, K., Beutin, L., Ghilardi, A. C. R., Gomes, T. A. T., Liberatore, A. M. A., Castro, A. F. P. & Elias, W. P. (2006). Enteroaggregative *Escherichia coli* from humans and animals differ in major phenotypical traits and virulence genes. *FEMS Microbiol Lett* **256**, 251–257.

Valle, G. R. F., Gomes, T. A. T., Irino, K. & Trabulsi, L. R. (1997). The traditional enteropathogenic *Escherichia coli* (EPEC) serogroup O125 comprises serotypes which are mainly associated with the category of enteroaggregative *E. coli*. *FEMS Microbiol Lett* **152**, 95–100.