ELSEVIER

FEMS Microbiology Letters 172 (1999) 145–151

FEMS MICROBIOLOGY LETTERS

Clonal relationship among invasive and non-invasive strains of enteroinvasive *Escherichia coli* serogroups

Marina B. Martinez [a,*], Thomas S. Whittan [b], Elizabeth A. McGraw [b], Josias Rodrigues [c], Luiz R. Trabulsi [d]

[a] *Faculdade de Ciencias Farmaceuticas, Departamento de Analises Clinicas e Toxicologicas, Universidade de Sao Paulo, Av. Prof. Lineu Prestes 580, Sao Paulo-SP, 05508-900 Brazil*

[b] *Institute of Molecular Evolutionary Genetics, Department of Biology, The Pennsylvania State University, University Park, PA 16802, USA*

[c] *Instituto de Biociencias-UNESP, Botucatu-SP, Brazil*

[d] *Instituto Butantan, Sao Paulo-SP, Brazil*

Received 3 November 1998; received in revised form 5 January 1999; accepted 5 January 1999

Abstract

The genetic relatedness among 96 invasive *Escherichia coli* belonging to several serogroups and 13 non-invasive of several serotypes that share the same O antigen was investigated by multilocus enzyme electrophoresis analysis. The invasive strains were isolated in different parts of the world and most of them recovered from dysentery. Twenty-nine electrophoretic types were distinguished and the most invasive strains were found to belong to two major lineages. These results suggested that the invasive ability in these strains has evolved in divergent chromosomal backgrounds, presumably through the horizontal spread of plasmid-borne invasion genes. The maintenance of invasive phenotypes in separate lineages suggests that this ability confers a selective advantage to invasive strains.

Keywords: Clonal relationship; Multilocus enzyme electrophoresis; Genetic polymorphism; Enteroinvasive *Escherichia coli*

1. Introduction

Some *Escherichia coli* strains are positive in the Serény test for invasiveness and cause a Shigella-like illness in humans, both children and adults [1–6]. The ability to invade colonic epithelial cells is mediated by a large plasmid encoding genetic determinants required for invasion [1,3]. These *E. coli* strains are known as enteroinvasive *E. coli* (EIEC) and this name has been universally accepted [7]. The EIEC corresponds to bioserotypes found in a dozen of *E. coli* serogroups [8]. Interestingly, some of these O antigens are identical or similar to *Shigella* O antigens [7].

The role played by EIEC in endemic diarrheal disease has not been investigated extensively. However, some studies indicated that these bacteria can be isolated with relatively high frequency depending on the population investigated. In the city of Sao Paulo, in Brazil for example, EIEC has been found in 5–7% of children living in medium-income families

* Corresponding author. Tel.: +55 (11) 818-3636; Fax: +55 (11) 813-2197; E-mail: mbmartin@usp.br

0378-1097/99/$20.00
PII: S0378-1097(99)00031-2

and in 20% of children who live in a slum, in the outskirts of the city. Both groups studied involved children older than 2-years-old, and they all showed acute diarrhea [8]. A similar frequency to that of Sao Paulo, has been reported by Echeverria et al. in Bangkok [9]. Outbreaks of food-borne infections due to EIEC have also been reported elsewhere [4].

Studies carried out on the biochemical characteristics of EIEC strains have shown that these bacteria do not decarboxylase lysine and except for a few serotypes, most strains are not motile. Strains that possess EIEC O antigens but decarboxylase lysine are always Sereny negative. These characteristics suggest that EIEC strains correspond to typical *E. coli* bioserotypes [10,11].

Multilocus enzyme electrophoresis (MLEE), which has long been a standard method in eukaryotic population genetics and systematics, has been used in studies to estimate the genetic diversity and structure in natural populations of a variety of species of bacteria. This research has established basic population genetic frameworks for the analysis of variation in serotypes and other phenotypic characters and has provided extensive data for systematics and useful marker systems for epidemiology [12].

In the present study, we determined the genetic relationship among invasive strains in an effort to learn more about the evolution of the invasive phenotype. We examined isolates representing ten EIEC serogroups commonly recovered from dysentery and characterized chromosomal backgrounds by MLEE. We compared the MLEE profile of EIEC strains with that of non-invasive, motile strains belonging to the same serogroups.

2. Materials and methods

2.1. Bacterial isolates

A total of 96 enteroinvasive and 13 non-enteroinvasive *E. coli* strains isolated from patients with diarrhea from different countries, between 1965 and 1988 were studied in this study. The strains from Brazil were collected from different cities. A total of 10 EIEC serotypes were studied, 98% were collected from cases of diarrhea in humans, 68% of the isolates were from Brazil, and remaining strains were isolated in different countries. The biochemical profiles and the serogroups of all strains were performed by standard methods [7]. Invasiveness was assessed by the Sereny test [13]. The number of invasive strains and their O and flagellar antigens of isolates are listed in Table 1.

2.2. Multilocus enzyme electrophoresis.

Genetic variation was analyzed by horizontal starch gel electrophoresis to detect polymorphisms for 20 enzyme loci, as previously described [14]. Electromorphs were compared to standard mobility markers and assigned mobility ranks by the rate of migration. Isolates lacking a particular enzyme activity were designated null for that particular locus. Each distinctive array of alleles was designated an electrophoretic type or ET [14]. The allelic arrays for the 20 loci defining ETs are listed in Table 2.

2.3. Phylogenetic analysis

Phylogenetic relationships between ETs were determined based on a matrix of genetic distances between all pairs constructed by comparisons of the allelic arrays. The neighbor-joining algorithm was used to constructed the dendrogram with the program MEGA [15,16]. The resulting phylogenetic tree produced is a representation of the genetic divergence observed in the different in the chromosomal backgrounds.

3. Results

3.1. Enzyme polymorphism electrophoretic groups and clonal groups

All 109 isolates were examined for allelic variation at 20 enzyme loci by MLEE. Fifteen of 20 enzymes loci were polymorphic, with an average of three alleles per locus. Comparison of the allele profiles of strains revealed 29 electrophoretic types (ETs), 19 of which included enteroinvasive strains (Tables 1 and 2). Among the total 29 ETs, 15 were represented by multiple isolates recovered from unrelated hosts at

Table 1
Electrophoretic type (ET), serotype, Séreny test and lysine decarboxylase of EIEC strains

ET	Serotype	No. of strains	Serény test	Lysine decarboxylase	Period of isolation	Country of origin
1	O28ac:H$^+$	3	−	+	1971−1981	Brazil/Japan
2	O28ac:H$^+$	1	−	+	1980	Japan
3	O136:H25	1	−	+	1980	Japan
4	O136:H40	1	−	+	1982	Brazil
5	O124:H19	1	−	+	1982	Brazil
6	O29:H21	1	−	+	1978	Brazil
7	O124:H7	2	+	+	1962	USA
	O136:NM	1	+	−	1985	Brazil
8	O29:H10	1	−	+	1978	Brazil
9	O29:H$^+$	1	−	+	1984	Mexico
9	O143:NM	1	+	−	1979	Brazil
10	O143:NM	1	+	−	1985	Brazil
11	O144:NM	5	+	−	1966–1984	Brazil/Japan
12	O136:NM	1	+	−	1977	Brazil
13	O124:NM	1	+	−	1980	French
14	O28ac:NM	10	+	−	1977–1985	Brazil/Chile
	O28ac:NM	3[a]	−	−		Japan/USA
15	O28ac:NM	4	+	−	1978–1982	Brazil/Chile
						Japan/USA
	NT	1	+	−	1981	USA
16	O164:NM	3	+	−	1980–1985	Brazil/Japan
17	O136:NM	11	+	−	1968–1983	Brazil/Japan
						Hungary
						Bangladesh
18	O29:NM	6	+	−	1976–1982	Brazil/Chile
						USA
19	O143:NM	6	+	−	1965–1982	Brazil/Chile
						Japan
20	O167:NM	5	+	−	1969–1985	Brazil
21	O136:H46	1	−	+	1969	Unknown
22	O152:NM	8	+	−	1964–1978	Brazil
23	O135:NM	3	+	−	1988	Brazil
24	O152:NM	2	+	−	1968	Brazil
	O152:NM	1	−	−		
25	O124:NM	16	+	−	1965–1988	Brazil/Japan
25	O124:H30	3	+	−		Hungary/USA
26	O164:NM	1	+	−	1985	Brazil
27	O124:NM	1[b]	−	−	1985	Brazil
28	0124:H19	1	−	+	1977	Brazil
29	O143:H$^+$	1	−	+	1979	Brazil

[a]One strain isolated from cheese.
[b]Isolated from water.

different times and different countries. Thus, these ETs were considered to mark widespread bacterial clones. In all but 4 cases, the isolates of the same ET belonged to the same serotype. The exceptions were: (1) ET 7, which comprised two invasive isolates of O124:H7 serotype and a non-motile O136:NM strain; (2) ET 9 included a non-invasive strain serotype O29 motile and an invasive O143:NM strain; (3) ET 15 included an O28:NM and an O non-typable strain, both of which were non-motile and invasive, and (4) ET 25, for which most of the O124 invasive isolates were non-motile,

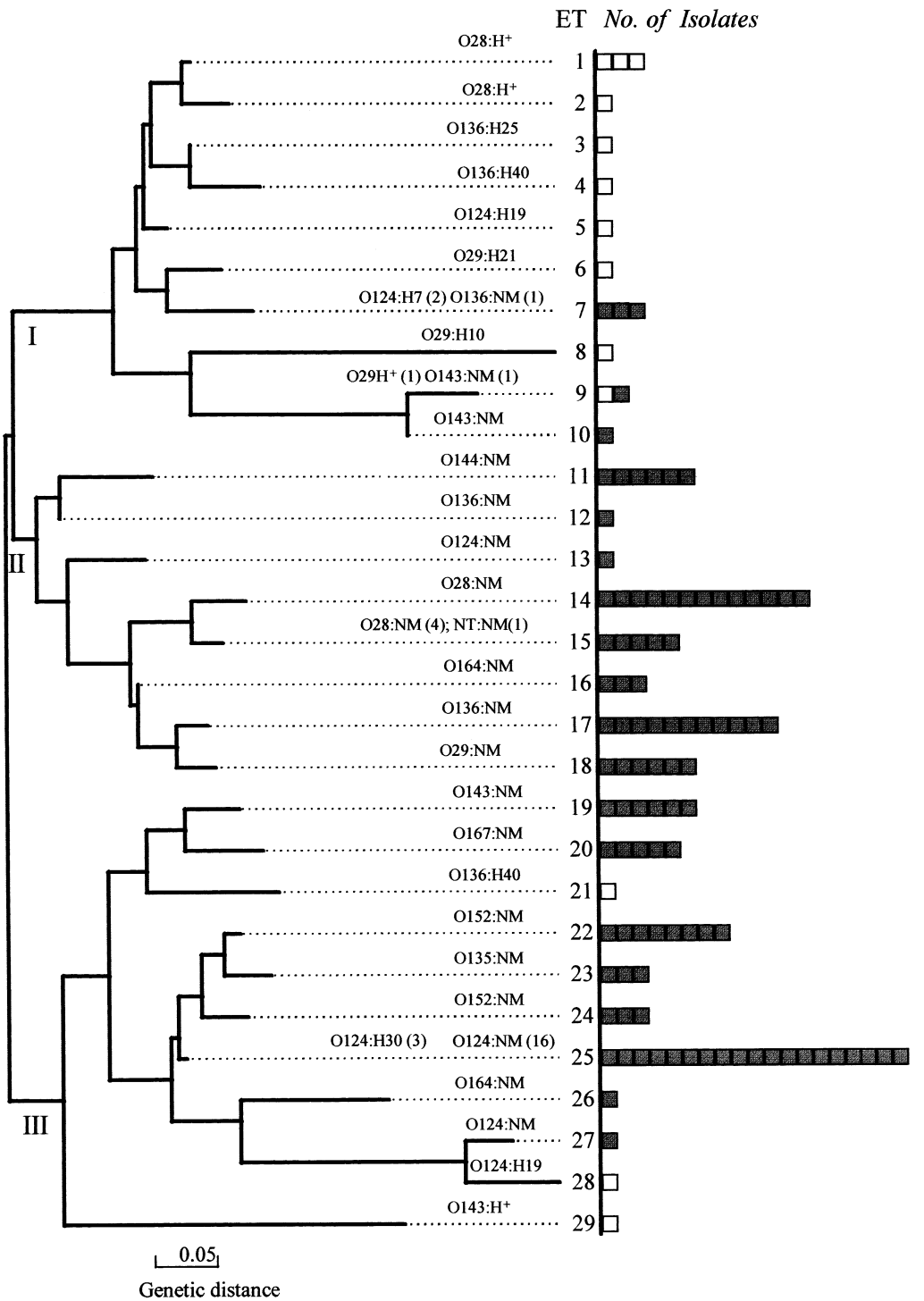

Fig. 1. Dendrogram of genetic relationship of 29 ETs of EIEC and non-EIEC strains. Genetic distance is measured in terms of detectable codon differences per enzyme locus. ■, EIEC strains; □, non-EIEC strains.

Table 2
Allele combination for 20 polymorphic enzymes that define 19 electrophoretic types (ETs) among enteroinvasive *E. coli* serotypes

ET	No. of strains	Serotype	PGI	IDH	ACO	G3P	PE2	AK	MDH	PGD	M1P	GOT	BGA	ADH	MPI	G6P	IPO	CAK	NSP	TDH	SKD	GLU
1	3	028:H+	6	5	6	4	5	2	4	62	8	6	5	6	102	4	4	4	2	4	8	2
2	1	O28:H+	6	5	6	4	5	2	4	62	8	6	6	0	102	4	4	4	2	4	8	2
3	1	O136:H25	6	5	6	4	5	2	4	6	8	6	5	0	102	4	4	4	2	4	8	2
4	1	O136:H40	6	5	6	4	5	2	4	6	8	4	5	6	102	4	4	4	2	4	8	2
5	1	O124:H19	6	5	6	4	5	2	4	8	8	6	5	0	102	4	4	4	2	4	8	2
6	1	O29:H21	6	5	6	4	5	1	4	13	8	6	5	0	102	4	4	4	2	4	8	2
7	2	O124:H7	6	5	6	4	5	1	4	8	8	6	0	6	102	4	4	4	2	4	8	2
	1	O136:NM	6	5	6	4	5	1	4	8	8	6	0	6	102	4	4	4	2	4	8	2
8	1	O29:H10	6	7	6	4	5	1	4	14	8	6	74	0	2	4	4	4	2	4	8	4
9	1	O29:H+	8	5	6	4	4	2	4	12	8	6	74	6	10	4	4	4	2	4	8	4
	1	O143:NM	8	5	6	4	4	2	4	12	8	6	74	6	10	4	4	4	2	4	8	4
10	1	0143:NM	8	5	6	4	4	2	4	12	8	6	74	6	10	4	4	4	2	4	8	2
11	5	O144:NM	5	5	82	4	5	2	4	6	8	6	5	4	8	4	4	4	2	4	8	2
12	1	O136:NM	5	5	6	4	5	2	4	6	8	6	0	0	8	4	4	4	2	4	8	2
13	1	O124:NM	5	5	82	4	5	2	4	6	8	6	0	4	8	4	4	4	2	4	8	2
14	13	O28ac:NM	5	5	6	4	5	2	4	62	6	6	9	4	8	4	4	4	1	4	8	2
15	4	O28ac:NM	5	5	6	4	5	2	4	62	8	6	9	4	8	4	4	4	1	4	8	2
15	1	O:NM	5	5	6	4	5	2	4	62	8	6	9	4	8	4	4	4	1	0	8	2
16	3	O164:NM	5	5	6	4	5	2	4	8	8	6	0	0	8	4	4	4	1	4	8	2
17	11	O136:NM	5	5	6	4	5	2	4	6	8	6	8	4	8	4	4	4	1	4	8	2
18	6	O29:NM	5	5	6	4	5	2	4	15	8	6	8	4	8	4	4	4	1	4	8	2
19	6	O143:NM	5	5	6	4	4	2	4	8	6	6	8	6	4	4	4	4	2	4	8	2
20	5	O167:NM	6	5	6	4	4	2	4	6	6	6	8	6	4	4	4	4	2	4	8	2
21	1	O136:H46	5	5	6	4	5	2	4	4	6	6	8	6	4	4	4	4	2	4	6	2
22	8	O152:NM	5	2	6	4	5	2	4	6	6	6	12	6	8	4	4	4	2	4	8	2
23	3	O135:NM	5	2	6	4	5	2	4	4	6	6	12	0	8	4	4	4	2	4	8	2
24	3	O152:NM	5	2	6	4	5	2	4	6	6	6	12	6	10	4	4	4	2	4	8	2
25	19	O124:H30/NM	5	2	6	4	5	2	4	8	6	6	0	6	8	4	4	4	2	4	8	2
26	1	O164:NM	5	2	7	4	5	2	4	8	6	6	8	6	8	2	4	4	2	4	8	2
27	1	O124:NM	5	2	7	4	7	2	4	6	6	6	8	0	2	4	4	4	2	4	4	2
28	1	O124:H19	5	2	7	4	7	2	4	6	7	6	8	4	4	4	4	4	2	4	4	2
29	1	O143:H+	5	5	5	4	23	2	4	8	4	6	5	0	10	4	4	4	2	4	7	2

PGI, glucosephosphate isomerase; IDH, isocitrate dehydrogenase; ACO, aconitase; G3P, glyceraldehyde-3-phosphate dehydrogenase; PE2, phenylalanyl-leucine peptidase; AK, adenylate kinase; MDH, malate dehydrogenase; PGD, gluconate-6-phosphate dehydrogenase; M1P, mannitol-1-phosphate dehydrogenase; GOT, aspartate aminotransferase; BGA, β-galactosidase; ADH, alcohol dehydrogenase; MPI, mannose-phosphate isomerase; G6P, glucose-6-phosphate dehydrogenase; IPO, indophenol oxidase; CAK, carbamate kinase; NSP, nucleoside phosphorylase; TDH, threonine dehydrogenase; SKD, shikimate dehydrogenase; GLU, glutamate dehydrogenase.

except for three isolates that expressed flagella antigen type H30.

3.2. Genetic analysis of the population

Genetic distance between each pair of the 29 ETs was estimated from the number of allelic differences and analyzed with the average linkage algorithm. The resulting dendrogram shows that EIEC strains fall into three distinct groups (labeled I–III) at a genetic distance greater than 0.05 (Fig. 1). Cluster I includes 78% of the non-invasive serotypes. Cluster II comprises the invasive serotypes O28:H$^-$, O29:H$^-$, O124:H$^-$, O136:H$^-$, O144:H$^-$ and O164:H$^-$. Cluster III comprises the invasive serotypes O124:H$^-$, O135:H$^-$, O143:H$^-$, O152:H$^-$, O167:H$^-$. The serotypes O124 and O164 are included in the both clusters (Fig. 1), but the most of

EIEC strains serotype O124:H$^-$ are in the cluster III (95%) and the serotype O164:H$^-$ the most of them are in cluster II (75%). Most of invasive strains belong to clusters II and III which included 45 (47%) and 46 (48%) of the total 96 invasive strains, respectively.

4. Discussion

Genetic analysis of the population of EIEC has shown most strains fall into two monophyletic clusters (II and III). This is inferred from genetic distances between chromosomal genotypes based on protein polymorphism detected by enzyme electrophoresis. The two EIEC clusters are not closely related to *E. coli* strains which possess other mechanism for causing diarrheal disease (data not shown).

The occurrence of the invasive ability in diverse backgrounds suggests that lateral transfer of plasmid-borne genes has been important in the evolution of this virulence phenotype [17]. Strains that lack the invasive phenotype, based on the Sereny test, and that are found within the clusters of invasive strains suggest a recent loss of the invasive property. The recent loss of the invasive phenotype could be due to either the loss or chromosomal integration of the invasive plasmid. Integration represses expression of invasion genes and may be advantageous when bacteria are outside of host cells or face adverse environmental conditions [18].

The present study showed 29 ETs, 19 of which included enteroinvasive *E. coli* strains. The similarity of ETs of EIEC strains having the same bioserotype, but varying in geographical origin. We observed that there was a high degree of divergence among invasive strains and non-invasive *E. coli* strains (Fig. 1). Most of the strains that did not belong to the major clusters II and III were negative for invasiveness were motile and decarboxylase lysine positive (Table 1). These results are in agreement with those found by Silva and co-workers. [10], who showed that invasive *E. coli* strains belong to bioserotypes that do not decarboxylate lysine and with a few exceptions, are not motile. Bando et al. [19] showed the same distinction between EIEC and non-EIEC strains by a genetic study based on random amplified polymorphic DNA (RAPD).

Acknowledgments

This work was supported by grants from FAPESP (94/3014-0).

References

[1] Fernandes, M.R. and Trabulsi, L.R. (1967) A new *Escherichia coli* serotype causing experimental keratoconjunctivitis in the guinea-pig (Culture 412 DEC-66). Preliminary report. Rev. Inst. Med. Trop. São Paulo 9, 62.
[2] Fernandes, M.R. and Trabulsi, L.R. (1969) Antigenic identity of culture 193T-64 and *Escherichia coli* O136:K78(B22). Rev. Inst. Med. Trop. São Paulo 11, 101–103.
[3] Formal, S.B. and Hornick, R.B. (1979) Invasive *Escherichia coli*. J. Infect. Dis. 137, 641–644.
[4] Gangarosa, E.J. (1978) Epidemiology of *Escherichia coli* in the United States. J. Infect. Dis. 137, 634–638.
[5] Levine, M.M. (1987) *Escherichia coli* that cause diarrhea: enterotoxigenic, enteropathogenic, enteroinvasive, enterohemorragic, and enteroadherent. J. Infect. Dis. 155, 377–389.
[6] Trabulsi, L.R., Fernandes, M.R. and Zuliani, M.E. (1967) Novas bactérias patogênicas para o intestino do homem. Rev. Inst. Med. Trop. São Paulo 9, 31–39.
[7] Ewing, W.H. (1986) Edwards and Ewing's identification of Enterobacteriaceae, 4th edn. Elsevier Science, New York, NY.
[8] Toledo, M.R.F. and Trabulsi, L.R. (1990) Frequency of enteroinvasive *Escherichia coli* in children with diarrhea and healthy controls in São Paulo, Brazil. Rev. Microbiol. São Paulo 21, 1–4.
[9] Echeverria, P., Sethabutr, O., Serichantalergs, O., Lexomboon, U. and Tamura, K. (1992) *Shigella* and enteroinvasive *Escherichia coli* infections in households of children with dysentery in Bangkok. J. Infect. Dis. 165, 144–147.
[10] Silva, R.M, Toledo, M.R.F. and Trabulsi, L.R. (1980) Biochemical and cultural characteristics of invasive *Escherichia coli*. J. Clin. Microbiol. 11, 441–444.
[11] Toledo, M.R.F. and Trabulsi, L.R. (1983) Correlation between biochemical and serological characteristics of *Escherichia coli* and results of the Serény test. J. Clin. Microbiol. 17, 419–421.
[12] Selander, R.K. and Musser, J.M. (1990) Population genetics of bacterial pathogenesis. In: The Bacteria, Vol. XI, pp. 11–36. Academic Press, New York, NY.
[13] Serény, B. (1955) Experimental *Shigella* keratoconjunctivitis. A preliminary report. Acta. Microbiol. Acad. Sci. Hung. 2, 293–296.
[14] Selander, R.K., Caugantu, D.A., Ochman, H., Musser, J.M., Gilmour, M.H. and Whittam, T.S. (1986) Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. Appl. Environ. Microbiol. 51, 873–884.
[15] Saitou, N. and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.
[16] Kumar, S., Tamura, K. and Nei, M. (1993) MEGA: molec-

ular evolutionary genetics analysis, version 1.0. The Pennsylvania State University, University Park.

[17] Whittam, T.S., Wolfe, M.L., Wachsmuth, I.K., Orskov, F., Orskov, I. and Wilson, R.A. (1993) Clonal relationship among *Escherichia coli* strains that cause hemorraghic colitis and infantile diarrhea. Infect. Immun. 61, 1619–1629.

[18] Zagaglia, C., Casalino, M., Colonna, B., Conti, C., Calconi, A. and Nicoletti, M. (1991) Virulence plasmids of enteroinvasive *Escherichia coli* and *Shigella flexneri* integrate into a specific site on the host chromosome: integration greatly reduces expression of plasmid-carried virulence genes. Infect. Immun. 59, 792–79.

[19] Bando, S.V., Valle, G.R.F., Martinez, M.B., Trabulsi, L.R. and Moreira-Filho, C.A. (1998) Characterization of enteroinvasive *Escherichia coli* and *Shigella* strains by RAPD analysis. FEMS Microbiol. Lett. 165, 159–165.