Genetic diversity of locus of enterocyte effacement genes of enteropathogenic Escherichia coli isolated from Peruvian children

C. A. Contreras,1 T. J. Ochoa,1,2 J. Ruiz,3,4 D. W. Lacher,5 D. Durand,1 C. DebRoy,6 C. F. Lanata7,8 and T. G. Cleary2

The aim of this study was to determine the frequency and allele associations of locus of enterocyte effacement encoded esp and tir genes among 181 enteropathogenic Escherichia coli (EPEC) strains (90 diarrhoea-associated and 91 controls) isolated from Peruvian children under 18 months of age. We analysed espA, espB, espD and tir alleles by PCR-RFLP. EPEC strains were isolated with higher frequency from healthy controls (91/424, 21.7 %) than from diarrhoeal samples (90/936, 9.6 %) (P, 0.001); 28.9 % of diarrhoeal and 17.6 % of control samples were typical EPEC (tEPEC). The distribution of espA alleles (alpha, beta, beta2 and gamma) and espD alleles (alpha, beta, gamma and a new variant, espD-N1) between tEPEC and atypical EPEC (aEPEC) was significantly different (P, 0.05). espD-alpha was more common among acute episodes (P, 0.05). espB typing resulted in five alleles (alpha, beta, gamma and two new sub-alleles, espB-alpha2 and espB-alpha3), while tir-beta and tir-gamma2 were the most common intimin receptor subtypes. Seventy-two combinations of espA, espB, espD and tir alleles were found; the most prevalent combination was espA-beta, espB-beta, espD-beta, tir-beta (34/181 strains), which was more frequent among tEPEC strains (P < 0.05). Our findings indicate that there is a high degree of heterogeneity among EPEC strains isolated from Peruvian children and that aEPEC and tEPEC variants cluster.

INTRODUCTION

Enteropathogenic Escherichia coli (EPEC) is an important group of diarrhoeal pathogens of young children living in developing countries (Chen & Frankel, 2005; Scaletsky et al., 2002). EPEC is characterized as typical or atypical depending on the presence of the EPEC adherence factor (EAF) virulence plasmid encoding bundle-forming pili, which are associated with a pattern of intestinal epithelial cell attachment known as localized adherence (Donnenberg et al., 1992). A key characteristic of all EPEC is a chromosomally located pathogenicity island named the locus of enterocyte effacement (LEE). Genes located on the LEE encode structural components of a type III secretion–translocation apparatus, factors enabling the bacterium to adhere intimately to intestinal epithelial cells (eae and tir).
secreted and effector proteins, chaperones and transcriptional regulators (Ler, GrlR/A) (Deng et al., 2004). The coordinated expression of these genes finally causes attaching-and-effacing lesions, histopathological lesions characterized by intimate bacterial adherence to the host cell plasma membrane, leading to destruction of the enterocyte microvilli and induction of cytoskeletal rearrangements beneath adherent bacteria (Frankel et al., 1998; Nataro & Kaper, 1998; Wong et al., 2011). The effector proteins are Tir, EspG, EspF, EspZ, Map and EspH; and the translocators are EspA, EspD and EspB, required for translocating the effectors into host cells (Dean & Kenny, 2009). Tir, EspB and EspD are transferred into the host cells through the translocation machinery formed by the type III secretion system. Translocated EspB and EspD are integrated into the cytoplasmic membrane of the target cells and form a pore that allows other molecules to enter these cells (Goffaux et al., 2001). Analysis of the LEE region shows that the core LEE is largely conserved, particularly among genes encoding the type III secretion system, whereas genes encoding effector proteins display a higher degree of variability (Müller et al., 2009). In addition to the high levels of sequence polymorphism observed in the eae gene (Blanco et al., 2006; Contreras et al., 2010; Lacher et al., 2006), allelic variants in the tir (alpha, beta, gamma and gamma2), espA (alpha, beta, beta2 and gamma), espB (alpha, beta and gamma) and espD (alpha, beta and gamma) genes have also been described (Afset et al., 2008; China et al., 1999; Garrido et al., 2006; Goffaux et al., 2001; Nielsen & Andersen, 2003; Yuste et al., 2008). In contrast, the esc (E. coli secreted components) and sep genes are more conserved (Goffaux et al., 2001). The differentiation of eae, tir and esp alleles is an important tool for EPEC typing as well as in epidemiological and clonal studies (Garrido et al., 2006; Yuste et al., 2008). Few studies have evaluated the associations between the tir, espA, espB and espD allele variants (Afset et al., 2008; Garrido et al., 2006). In the present study, we classified EPEC strains based on associations of esp and tir alleles using PCR-RFLP and investigated the relationships between these esp–tir associations and the characteristics of the diarrhoeal episodes.

**METHODS**

**Strains.** The 181 EPEC strains examined in this study were isolated from a cohort epidemiological study of diarrhoea in infants from 2 to 18 months of age (Contreras et al., 2010; Ochoa et al., 2009). In the study, 1360 stools samples were obtained from 1034 children; 936 samples were from children with diarrhoea (case patients) and 424 were from children without diarrhoea (controls). Five lactose-positive colonies were analysed by real-time PCR to detect diarrhoeagenic E. coli as described previously (Guion et al., 2008). One hundred and twenty of the 181 strains were analysed previously for detecting variability in eae, bfpA and perA genes and the association with clinical characteristics (Contreras et al., 2010).

**Clinical data and serotyping.** Clinical information on the diarrhoeal episodes obtained during the cohort study was analysed by a modified Vesikari score (Ruuska & Vesikari, 1990) to quantify the severity of the episode. The score included: duration of diarrhoea in days (0–3 points), maximum number of stools per day during the episode (1–3), number of days with vomiting (0–3), maximum number of emesis per day during the episode (0–3), presence of fever (0–1), dehydration (0–3) and treatment (0–2). The maximum possible score was 18. A ‘mild’ score was 0–8 points, a ‘moderate’ score was 9–14 points and a ‘severe’ score was 15–18 points. In the analysis, moderate and severe scores were grouped together because of the low number (n=3) of patients with a severe score.

The serotyping was performed at the E. coli Reference Center at the Pennsylvania State University according to standard methods (Contreras et al., 2010).

**Detection of virulence genes.** EPEC isolates were examined for bfpA, espA, espB, espD and tir genes by PCR using the primers and conditions listed in Table 1. PCR for the five genes was performed in a 25 μl reaction mixture containing 2.5 μl 2.5 mM of each dNTP (Promega), 2.5 μl Taq polymerase (Kappa Biosystems), 0.5 μl MgCl2 (Kappa Biosystems) and 2 μl DNA template. PCR amplification was performed in a thermocycler (iCycler; Bio-Rad) under the conditions listed in Garrido et al. (2006), Guion et al. (2008) and Lacher et al. (2006) (also see Table 1 available in JMM Online). The mixture was held at 72 °C for 7 min after the final cycle before cooling at −20 °C. Amplified products were analysed by using 1.5% agarose gel electrophoresis and visualized by staining with ethidium bromide. PCR products were analysed by RFLP. The positive PCR control was the EPEC strain E2348/69 and the negative control was E. coli C600.

**PCR-RFLP.** The in silico RFLP tool of the EcMLST website (http://www.shigatox.net/ecmlst/cgi-bin/insilicorflp) was used to predict the restriction patterns of all the genes except bfpA (Table 1). One restriction enzyme digest was used for all genes analysed except for

**Table 1. Expected RFLPs of esp PCR amplicons**

The digestion patterns are shown (bp).

<table>
<thead>
<tr>
<th>espA allele</th>
<th>AluI</th>
<th>Ddel</th>
<th>espB allele</th>
<th>BstUI</th>
<th>espD allele</th>
<th>BstUI</th>
<th>tir allele</th>
<th>Ddel</th>
</tr>
</thead>
<tbody>
<tr>
<td>alpha</td>
<td>237</td>
<td>173, 64</td>
<td>alpha</td>
<td>444, 147, 135</td>
<td>alpha</td>
<td>335, 335</td>
<td>alpha</td>
<td>201, 157, 90, 85</td>
</tr>
<tr>
<td>beta</td>
<td>175, 62</td>
<td>143, 64, 34</td>
<td>beta</td>
<td>523, 135, 56</td>
<td>beta</td>
<td>467, 203</td>
<td>beta</td>
<td>530</td>
</tr>
<tr>
<td>beta2</td>
<td>103, 72, 62</td>
<td>173, 64</td>
<td>gamma</td>
<td>523, 78, 56, 45</td>
<td>gamma</td>
<td>377, 275</td>
<td>gamma2</td>
<td>236, 192, 99</td>
</tr>
<tr>
<td>gamma</td>
<td>175, 62</td>
<td>207, 30</td>
<td>espB-alpha2</td>
<td>444, 279</td>
<td>espD-N1</td>
<td>670</td>
<td>tir-alpha2</td>
<td>376, 157</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>espB-alpha3</td>
<td>714</td>
<td></td>
<td></td>
<td>tir-N2</td>
<td>291, 147, 89</td>
</tr>
</tbody>
</table>

http://jmm.sgmjournals.org
espB (digested with two separate restriction enzymes). Digestion was performed in 20 μl reaction mixtures at a final concentration of 10 U enzyme (New England Biolabs), 1 x reaction buffer, 15.0 μl unpurified PCR product and distilled water to complete the final volume. The samples were incubated overnight at the temperature indicated by each enzyme manufacturer. After incubation, 15 μl each digest was visualized on ethidium bromide-stained 3.0 % agarose gels by illumination with UV light.

Sequencing of the esp and tir genes. The strains that shared an undefined RFLP pattern were sequenced and analysed to establish the genetic relationship with known alleles. PCR products were purified with the QiAquick PCR Purification kit (Qiagen). Sequencing was performed by Macrogen (Korea) using an automatic DNA sequencer (Applied Biosystems 3730XL) and used for phylogenetic analyses. DNA sequences of esp genes were edited with BioEdit v4.8.10 (http://www.mbio.ncsu.edu/bioedit/bioedit.html).

Statistical analysis. The allelic frequencies obtained in each group and the allelic distributions in each population were compared using the program GenAlEx 6.3 and GenPop on the web tool (http://genepop.curtin.edu.au/). The comparisons between groups were made using the chi-square or Fisher’s exact test. P-values <0.05 were considered significant.

Nucleotide sequence accession numbers. The GenBank accession numbers for the sequences of the new espB (espB-alpha3 and espB-alpha2), espD (espD-N1) and tir (tir-alpha2) alleles are JN571730, JN571731, JN571732 and JN571733, respectively.

RESULTS

Prevalence of esp, tir and bfpA genes

A total of 181 E. coli strains were classified as EPEC based on the presence of cac and absence of stx1/stx2 by PCR. EPEC strains were isolated with higher frequency from healthy controls (91/424, 21.5 %) than from diarrhoea samples (90/936, 9.6 %) (P<0.0001). Based on the presence of bfpA, 42/181 (23.2 %) strains were classified as tEPEC [26/90 (28.9 %) diarrhoea samples and 16/91 (17.6 %) controls, P=0.08].

The espA gene was present in 88/90 (97.8 %) diarrhoea samples and 89/91 (97.8 %) controls (Table 2). RFLP typing resulted in four previously reported espA alleles: espA-alpha [48 strains, 5/42 (11.9 %) in tEPEC vs 43/139 (30.9 %) in aEPEC; P<0.05], espA-beta (57 strains), espA-beta2 (64 strains) and espA-gamma (8 strains). The distribution of these alleles between diarrhoea samples and controls was similar (Table 2), although the difference in the distribution of these alleles among tEPEC and aEPEC was marginally significant.

For espB, the primers failed to amplify in 11 strains. Of the espB-positive strains (170/181, 93.9 %), typing resulted in three previously reported espB alleles, espB-alpha (53 strains), espB-beta (52 strains) and espB-gamma (15 strains), and two new sub-alleles designated espB-alpha2 (41 strains) and espB-alpha3 (9 strains). The distribution of these alleles was similar between diarrhoea samples and controls, and between tEPEC and aEPEC strains (Table 2).

The espD-positive strains were subtyped as espD-beta (64 strains), espD-N1 (40 strains), espD-gamma (28 strains) and espD-alpha (19 strains). A high number of strains were espD-negative (30/181, 16.6 %). There was no difference in the distribution of espD alleles among diarrhoea samples versus controls, while the distribution of espD alleles was different between tEPEC vs aEPEC, specifically espD-N1 was more common in aEPEC (P<0.05).

The tir subtyping identified tir-beta (52/181, 28.7 %) and tir-gamma2 [6/42 (14.3 %) in tEPEC vs 45/139 (32.4 %) in aEPEC; P<0.05] as the most common intimin receptor subtypes, followed by tir-alpha [14/42 (33.3 %) in tEPEC vs 22/139 (15.8 %) in aEPEC; P<0.05], tir-alpha2 (21/181, 11.6 %) and tir-gamma (7/181, 3.9 %). Fourteen strains showed negative PCR for tir with the primers used in this study (Table 2).

Combinations among alleles of LEE genes

Seventy-two combinations of variants of the LEE genes espA, espB, espD and tir were present among the 181 EPEC strains (GenAlex) (Tables S2 and S3). Two of these combinations were more prevalent: espA-beta, espB-beta, espD-beta, tir-beta (34/181, 18.8 %); and espA-alpha, espB-alpha, espD-N1, tir-gamma2 (21/181, 11.6 %). The first LEE gene combination was associated with tEPEC (P<0.05). Overall, the distribution of the combinations of all alleles described (espA, espB, espD and tir) was different between tEPEC and aEPEC using GenPop (http://genepop.curtin.edu.au/) (P<0.05) (Table S2).

Analysis of sequenced LEE genes

The RFLPs were designed based on the allele sequences available in GenBank for all LEE genes (espA, espB, espD and tir) (Table S2). However, many strains did not show a previously described RFLP pattern assigned to the known alleles. The two espB, one espD, and one tir nontypable RFLP patterns were designated espB-alpha2, espB-alpha3, espD-N1 and tir-alpha2, respectively. The PCR products of the strains with these RFLP patterns were sequenced. Analysis of the sequenced regions of the espB-alpha2 and espD-N1 amplics revealed them to be identical to the espB and espD sequences found in the complete genome of O111: H− strain 11128. The espB-alpha3 allele was most similar to the espB allele of strain 11128 with 18 polymorphic sites within the sequenced region of the amplicon. The tir-alpha2 allele was most similar to the allele found in an O119:H9 strain (GenBank accession no. AJ633129) with 2 nt substitutions. A new allele found in an ONT:H19 strain named tir-N2 (GenBank accession number AB288104) was also identified.

Serotypes

One hundred and twenty of 181 EPEC strains were previously serotyped (Contreras et al., 2010). An O serogroup was identified in 105/181 (58 %) of the strains.
Table 2. Distribution of *esp* alleles among tEPEC and aEPEC strains and among diarrhoea and control samples

<table>
<thead>
<tr>
<th>Genes</th>
<th>Allele</th>
<th>Diarrhoea [n (%)] (n=90)</th>
<th>Control [n (%)] (n=91)</th>
<th>tEPEC [n (%)] (n=42)</th>
<th>aEPEC [n (%)] (n=139)</th>
<th>All EPEC [n (%)] (n=181)</th>
</tr>
</thead>
<tbody>
<tr>
<td>espA alleles</td>
<td>espA-alpha</td>
<td>24 (26.7)</td>
<td>24 (26.4)</td>
<td>5 (11.9)</td>
<td>43 (30.9)*</td>
<td>48 (26.1)</td>
</tr>
<tr>
<td></td>
<td>espA-beta</td>
<td>30 (33.3)</td>
<td>27 (29.7)</td>
<td>15 (35.7)</td>
<td>42 (30.2)</td>
<td>57 (31.5)</td>
</tr>
<tr>
<td></td>
<td>espA-beta2</td>
<td>33 (36.7)</td>
<td>31 (34.1)</td>
<td>22 (52.4)</td>
<td>42 (30.2)</td>
<td>64 (35.4)</td>
</tr>
<tr>
<td></td>
<td>espA-gamma</td>
<td>1 (1.1)</td>
<td>7 (7.7)</td>
<td>0 (0.0)</td>
<td>8 (5.8)</td>
<td>8 (4.4)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>2 (2.2)</td>
<td>2 (2.2)</td>
<td>0 (0.0)</td>
<td>4 (2.9)</td>
<td>4 (2.2)</td>
</tr>
<tr>
<td>espB alleles</td>
<td>espB-alpha</td>
<td>26 (26.7)</td>
<td>27 (29.7)</td>
<td>17 (40.5)</td>
<td>36 (25.9)</td>
<td>53 (29.3)</td>
</tr>
<tr>
<td></td>
<td>espB-beta</td>
<td>31 (34.4)</td>
<td>21 (23.1)</td>
<td>14 (33.3)</td>
<td>38 (27.3)</td>
<td>52 (28.7)</td>
</tr>
<tr>
<td></td>
<td>espB-gamma</td>
<td>8 (8.9)</td>
<td>9 (9.9)</td>
<td>2 (4.8)</td>
<td>13 (9.4)</td>
<td>15 (8.3)</td>
</tr>
<tr>
<td></td>
<td>espB-alpha2</td>
<td>16 (17.8)</td>
<td>25 (27.5)</td>
<td>7 (16.7)</td>
<td>34 (24.5)</td>
<td>41 (22.7)</td>
</tr>
<tr>
<td></td>
<td>espB-alpha3</td>
<td>5 (5.6)</td>
<td>4 (4.4)</td>
<td>1 (2.4)</td>
<td>8 (5.8)</td>
<td>9 (4.9)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>6 (6.7)</td>
<td>5 (5.5)</td>
<td>1 (2.4)</td>
<td>10 (7.2)</td>
<td>11 (6.1)</td>
</tr>
<tr>
<td>espD alleles</td>
<td>espD-alpha</td>
<td>7 (7.8)</td>
<td>12 (13.2)</td>
<td>7 (16.7)</td>
<td>12 (8.6)</td>
<td>19 (10.5)</td>
</tr>
<tr>
<td></td>
<td>espD-beta</td>
<td>38 (42.2)</td>
<td>26 (28.6)</td>
<td>19 (45.2)</td>
<td>45 (32.4)</td>
<td>64 (35.4)</td>
</tr>
<tr>
<td></td>
<td>espD-gamma</td>
<td>13 (14.4)</td>
<td>15 (16.5)</td>
<td>3 (7.1)</td>
<td>25 (18.0)</td>
<td>28 (15.5)</td>
</tr>
<tr>
<td></td>
<td>espD-N1</td>
<td>18 (20)</td>
<td>22 (24.2)</td>
<td>4 (9.5)</td>
<td>36 (25.9)*</td>
<td>40 (22.1)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>14 (15.6)</td>
<td>18 (19.8)</td>
<td>9 (21.4)</td>
<td>21 (15.1)</td>
<td>30 (16.6)</td>
</tr>
<tr>
<td>tir alleles</td>
<td>tir-alpha</td>
<td>18 (20)</td>
<td>18 (19.8)</td>
<td>14 (33.3)*</td>
<td>22 (15.8)</td>
<td>36 (19.9)</td>
</tr>
<tr>
<td></td>
<td>tir-beta</td>
<td>33 (36.7)</td>
<td>19 (20.9)</td>
<td>13 (31.0)</td>
<td>39 (28.1)</td>
<td>52 (28.7)</td>
</tr>
<tr>
<td></td>
<td>tir-gamma</td>
<td>2 (2.2)</td>
<td>5 (5.5)</td>
<td>1 (2.4)</td>
<td>6 (4.3)</td>
<td>7 (3.9)</td>
</tr>
<tr>
<td></td>
<td>tir-gamma2</td>
<td>21 (23.3)</td>
<td>30 (33.0)</td>
<td>6 (14.3)</td>
<td>45 (32.4)*</td>
<td>51 (28.2)</td>
</tr>
<tr>
<td></td>
<td>tir-alpha2</td>
<td>8 (8.9)</td>
<td>13 (14.3)</td>
<td>7 (16.7)</td>
<td>14 (10.1)</td>
<td>21 (11.6)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>8 (8.9)</td>
<td>6 (6.6)</td>
<td>1 (2.4)</td>
<td>13 (9.4)</td>
<td>14 (7.7)</td>
</tr>
</tbody>
</table>

*P<0.05 for the comparison of allele distribution between tEPEC and aEPEC.*
investigated. The remaining strains were nontypable with the O antisera used (76 strains). Forty-one strains belonged to classical EPEC serogroups (Trabulsi et al., 2002). The other nontypable strains belonged to 36 different serogroups. The most common O serogroups were O55 (14 strains), O111 (7 strains) and O119 (6 strains). Thirty different H-types were detected; the most common flagellar types were H27 (20 strains), H8 (16 strains), H7 (14 strains) and H19 (13 strains). Six strains were non-motile and the H-type was not identified in 11 strains. No single serotype was found in more than 3% of the strains investigated.

**Association with clinical characteristics**

To determine whether clinical characteristics of the diarrhoea episodes might be related to specific combinations of esp and tir alleles, single EPEC infections (without other pathogens, e.g., Shigella, Campylobacter, etc.) were evaluated (60 EPEC strains from diarrhoea). Disease severity, as suggested by the Vesikari score, was similar in the diarrhoea episodes associated with the different allele combinations (data not shown). Among all the diarrhoeal episodes (unique and coinfections), 56 cases had a duration of <7 days, while 34 cases lasted ≥7 days. There were no clear differences in the distribution of espA, espB or tir alleles among diarrhoea episodes of longer duration (≥7 days) and shorter episodes (data not shown). For espD, the allele distribution was significantly different in relation to the duration of the episode. Specifically, espD-alpha was more common in episodes of <7 days compared with episodes ≥7 days (12.5% vs 0%, respectively; P<0.05).

**DISCUSSION**

The differentiation of tir, espA, espB and espD alleles of the LEE pathogenicity island represents a potential tool for EPEC typing in pathogenesis, epidemiological and immunological studies. In this work, we defined a PCR-RFLP for typing of four alleles of espA, five alleles of espB, four alleles of espD and six alleles of tir.

aEPEC (139/181, 76.8%) constituted the majority of eae-positive isolates in this study and was common both in patients with diarrhoea (64/90, 71.1%) and in controls (75/91, 82.4%), as has been recently reported in developing and developed countries (Ochoa et al., 2008).

In the present study, we identified four variants for espA; we found that espA-beta2 was the more frequent allele, whereas in other studies espA-alpha or espA-beta were more frequent variants (Afset et al., 2006; Garrido et al., 2006; Goffaux et al., 2001). As reported by Garrido et al. (2006), we found that espB-alpha was the most frequent allele; espB-beta was the second most frequent allele in this study. In other studies, espB-beta was the most frequent allele detected, but the strains were principally isolated from animals (Afset et al., 2008; Goffaux et al., 2001; Yuste et al., 2008). For espD, the most frequent allele was espD-beta (35.4%) (Yuste et al., 2008), in contrast to other studies where most strains were not typable for espD (Garrido et al., 2006). For the tir alleles, tir-beta (28.7%) (Goffaux et al., 2001; Yuste et al., 2008) was the most common tir variant, in contrast to a previous study that found tir-alpha to be the most frequent (~44%) (Garrido et al., 2006). These differences in the prevalence of different alleles/sub-alleles may be related to the different geographical origin of the samples. In addition to the high variability observed in these genes, some EPEC strains were PCR-negative for the genes evaluated in this study (Table 2). These PCR-negative results may be due to the absence of the genes in these strains, or mainly because the primers used do not hybridize to the template DNA due to the high sequence variability of these genes in these specific strains.

Although no homologous associations were observed between eae and bfpA/perA alleles in these strains (Conteras et al., 2010), associations of the results of tir and eae alleles (located in the fifth polycistronic operon termed LEE5) (Deng et al., 2001) showed that 16% of the EPEC strains were eae-beta and tir-beta and 23% were eae-theta and tir-gamma2. These two associations were found principally in aEPEC (data not shown). This association between tir and eae alleles was maintained in association with the esp genes. Twenty-seven of 34 (79%) of the strains with associations espA-beta, espB-beta, espD-beta, tir-beta were eae-beta and tir-beta and 14 of 21 (67%) strains with associations espA-alpha, espB-alpha2, espD-N1, tir-gamma2 were eae-theta. These results were similar to those obtained in isolates from animals (Yuste et al., 2008). The associations of the tir and esp genes in EPEC strains were associated with variants of the eae gene but not with the origin of the strains.

We found a high number of LEE allele combinations compared to reports from other authors (Afset et al., 2008; China et al., 1999; Goffaux et al., 2001; Nielsen & Andersen, 2003; Yuste et al., 2008). China et al. (1999) and Goffaux et al. (2001) found only four LEE profiles in human and bovine attaching and effacing E. coli strains, Nielsen & Andersen (2003) identified seven LEE gene combinations in verocytotoxin-producing E. coli calf strains (20 strains) and Garrido et al. (2006) reported 12 combinations of these LEE genes among 25 human and animal Shiga toxin-producing E. coli and EPEC strains studied. The combination espA-beta, espB-beta, espD-beta, tir-beta was the most frequent (18.8%); it has been noted in reports of strains isolated from animal and human sources (China et al., 1999; Garrido et al., 2006; Goffaux et al., 2001; Yuste et al., 2008) (Table S3). The most frequent homologous association (espA-beta, espB-beta, espD-beta, tir-beta) of the esp genes was closely related to eae-beta (28/34, 82%), similar to the results obtained in other studies (Goffaux et al., 2001; Yuste et al., 2008). The second most frequent association observed was espA-alpha, espB-alpha2, espD-N1, tir-gamma2 (Table S3) and was associated with eae-theta. Most of the esp and tir associations were heterologous (alleles in combination with more than just one variant of the other LEE genes), consistent with the findings of Afset et al. (2008). They suggest that these different combinations within the LEE may be due to
horizontal exchange of smaller or entire LEE sequences between strains (Afset et al., 2008; Castillo et al., 2005). Variants of the LEE genes have been associated with the characteristics of colonization and tissue tropism (Torres et al., 2005). Such differences might influence the ability of the bacteria to induce diarrhoea. However, except for a marginally significant association of espD-alpha with diarrhoea episodes 7 days, none of the LEE gene variants were significantly associated with diarrhoea. There were some differences between tEPEC and aEPEC in the distributions of the alleles, principally in espA-alpha (related to tEPEC) and espD-alpha (related to aEPEC). tir-gamma2 (related to aEPEC), espD-N1 (related to aEPEC), and tir-alpha (related to tEPEC) were significantly associated with diarrhoea. There were some differences in the distributions of the alleles, particularly in espA-alpha (related to tEPEC) and espD-alpha (related to aEPEC). These differences between tEPEC and aEPEC in the genetic profiles of the isolates suggest that the strains isolated from diarrhoea cases are not identical to the strains isolated from controls. The large variability in allelic types suggests that this could be used for evaluation of outbreaks to determine the better characterization of the strains isolated. However, it is clear from this study that allergic reactions are not usually virulent.

**ACKNOWLEDGEMENTS**

This work was partially funded by the National Institutes of Health, USA, Public Health Service awards 1K01TW007405 (T.J.O) and R01-HD051716 (T.G.C); Agencia Española de Cooperación Internacional (AECID), Spain, Programa de Cooperación Interuniversitaria e Investigación Científica con Iberoamérica (D/019499/08, D/024648/09 and D/030509/10); and Fondo de Investigaciones Sanitarias, Spain. J.R. has a Miguel Servet fellowship (CP05/0130).

**REFERENCES**


