Molecular epidemiology of *Shigella flexneri* in a diarrhoea-endemic area of Lima, Peru

C. M. FERNANDEZ-PRADA1*, M. M. VENKATESAN1, A. A. FRANCO2, C. F. LANATA3, R. B. SACK2, A. B. HARTMAN1 AND W. SPIRA2

1 Walter Reed Army Institute of Research (WRAIR), Silver Spring, MD 20910, USA
2 Johns Hopkins University, Baltimore, MD 21205, USA
3 Instituto de Investigacion Nutricional, Lima, Peru

(Accepted 30 September 2003)

SUMMARY

A year-long community-based study of diarrhoeal diseases was conducted in Canto Grande, a periurban community in Lima, Peru. In 109 (34%) houses out of 323 that were visited, at least one individual was detected with shigellosis. The frequency of the 161 shigella isolates obtained was as follows: 117 *S. flexneri* (73%), 21 *S. boydii* (13%), 15 *S. dysenteriae* (9%), and 8 *S. sonnei* (5%). Using a non-radioactive *ipaH* gene probe as a molecular epidemiological tool, a total of 41 *S. flexneri* strains were shown to be distributed in 25 intra-family comparisons by pairs (icp). Further subdivision, based on a comparison of the serotype, plasmid profile, antibiotic resistances and *ipaH* hybridization patterns indicated that Group I, with 11 icp (44%), had strains that were identical, Group II with 8 icp (32%), had strains that were different and Group III with 6 icp (24%), had strains with the same serotype and identical *ipaH* profiles but with differences in other markers. This data indicates that a diversity of shigella clones circulated in this community resulting from both clonal spread and horizontal transfer of genetic elements. Furthermore, *ipaH* profiling of isolates can be used not only to differentiate between closely related shigella strains but also with other parameters, help to understand the dynamics of the generation of new clones of pathogenic bacteria.

INTRODUCTION

Each year, more than 5 million children die in developing countries due to diarrhoeal diseases, and approximately 10–20% of these deaths are caused by bacillary dysentery produced principally by *Shigella* [1–3]. A recent study estimated that the annual number of shigella episodes throughout the world was 164·7 million, of which 163·2 million were in developing countries with 1·1 million deaths [4]. Shigellosis is also a public-health problem in developed countries, but is restricted to closed groups such as the military, mental hospitals, day-care nurseries, prisons, or cruise-ships [2, 3]. In the United States more than 20000 cases were reported to the Centers for Disease Control, Atlanta, in 2001 [5].

In developing countries, shigellosis infections are caused principally by *S. flexneri*, followed by *S. sonnei*; while in developed countries, they are caused principally by *S. sonnei*, followed by *S. flexneri* [4, 6]. Contaminated food, flies, unclean water as well as the faecal–oral route have contributed to the transmission and high incidence of shigella organisms, especially for the small inoculum size required for infection
The 50% infectious dose (ID$_{50}$) of shigella strains for humans has been reported to be as low as 100–200 bacteria [7]. This explains why the disease can be easily spread by person-to-person contact as well as by contaminated food and water [3].

Shigella is not only pathogenic to humans, but also to non-human primates like gorillas, chimpanzees and monkeys. All virulent Shigella species and enteroinvasive Escherichia coli (EIEC) contain plasmids of 120–140 MDa, carrying genes that, in conjunction with genes present in ‘pathogenicity islands’ on the chromosome, allow these organisms to invade, multiply and spread from cell to cell to extend the foci of infection [2, 8]. Infection with virulent Shigella or EIEC results in the production of immune serum with antibodies to surface-located and secreted proteins of the bacteria [2, 9]. Among these, IpaA, IpaB, IpaC, IpaD, VirG and IpaH proteins are encoded by genes on the virulence plasmid and play a role in the pathogenesis of the disease.

The $ipaH$ gene is unique in that multiple copies have been detected on the invasion plasmid and on the chromosome of all Shigella serotypes and EIEC [8, 10–12], but not on the chromosome of non-pathogenic E. coli strains [13]. $ipaH$ was initially isolated from a λgt11 expression library of the invasion plasmid (pWR100) of S. flexneri 5, strain M90T [14]. The first sequence that was analysed ($ipaH_{7,8}$) did not have a HindIII site within its structural gene [11]. When this sequence was used as a probe against HindIII-digested DNA on Southern blots of M90T, five copies were identified on pWR100, designated $ipaH_{4,5}$, $ipaH_{4,6}$, $ipaH_{4,7}$, $ipaH_{4,8}$ and $ipaH_{4,9}$, with the subscripts indicating the sizes of the HindIII band which hybridized to the $ipaH_{7,8}$ probe [11]. The use of oligonucleotide probes of the $ipaH_{7,8}$ gene seems to be useful in genetic lineage and epidemiological studies of shigellae and EIEC [10, 15, 16] (Hartman AB, unpublished observations). The development of specific non-radioactive $ipaH$-related DNA probes to identify Shigella and EIEC has enabled investigators to determine the prevalence of these organisms in different populations, to identify mixed infection and rare serotypes, and to screen for Shigella and EIEC in large numbers of patients [16–18]. Additionally, the sensitivity and specificity of the polymerase chain reaction (PCR) for detection of DNA sequences specific to shigellae and EIEC, like the $ipaH$, virF and ial loci [15, 17–20] have indicated the importance of this technique in diagnostic and epidemiological shigella studies.

In the present study, a non-radioactive $ipaH$ probe was evaluated as a molecular epidemiological tool to detect genetic polymorphisms between S. flexneri strains from an endemic area of shigellosis in Lima, Peru. This procedure has been coupled to several other epidemiological markers such as serotype, biochemical reactions, antibiotic susceptibility and plasmid profile.

**MATERIALS AND METHODS**

**Epidemiological investigations**

As part of a prospective, community-based study of diarrhoeal diseases conducted in Canto Grande, a periurban community of low socioeconomic conditions and poor hygienic conditions in Lima, Peru [21–23], all children $<3$ years of age from a representative sample of 400 households were visited twice per week at their households by trained field workers as described elsewhere [24, 25]. A stool specimen or a rectal swab was collected from any child who developed diarrhoea (defined as three of more liquid or semi-liquid stools passed in a 24-h period or with one bloody stool) and were placed in a clean cup with a lid or in Cary–Blair (BBL) and transported with cold packs to the laboratory, where they were processed within 24 h of collection [25]. Between February 1985 and February 1986, 37 index-children $<2$ years of age with diarrhoea and a shigella strain identified in their stool sample were entered into this study. After obtaining informed consent, all family members from the household of the index-child were asked to participate in this study, as well as all family members of up to nine households selected from the same neighbourhood of the index-child. The selection was made applying the following procedures. Once a shigella strain was identified in the laboratory, a study team visited the household of the index-child as soon as possible ($<60$ h from the time the index-child was cultured) and all available family members were interviewed. Three households with a family member who had contact with the index-child since his/her diarrhoea started were identified, as well as three other households with a family member who had been in contact with a family member of the index-child, but not with the index-child. Three additional households from the same neighborhood of the index-child family without any type of recent contact with any member of the index-child were identified by door-to-door survey, starting in the same block of households.
of the index-child. The group of the index-child household and the three types of household controls constituted one family study of up to ten households each. A daily rectal swab was taken from all consenting individuals of any age from all these households during one week.

**Bacteriological investigations**

The rectal swab samples were directly inoculated in the field onto Salmonellla–Shigella, MacConkey, or Heektoen agar plates (Difco Laboratories, Detroit, MI, USA). The plates were maintained at room temperature and immediately transported to the Microbiology Laboratory of the Universidad Peruana Cayetano Heredia, Lima, Peru. The plates were incubated overnight at 37 °C. Non-lactose-fermenting colonies were inoculated onto tryptic sugar iron (TSI), lysine iron agar (LIA), and urea tubes and examined for motility, glucose fermentation, production of gas, indole, hydrogen sulphide, and hydrolysis of urea. Final identification was performed by agglutination with specific polyvalent antisera [Laboratorio Nacional de Referencia de Enteropatogenos (LANARE), Lima, Peru]. A total of 161 shigella strains were isolated from humans during the year of study. Only one colony per individual was immediately stored at −70 °C in brain heart infusion broth (BHI; Difco Laboratories) with 25% (v/v) glycerol. When needed, the stock cultures of bacteria were plated onto tryptic soy agar (TSA) plates supplemented with 0.05% Congo Red dye and 0.2% galactose and incubated overnight at 37 °C. Congo Red-positive (CR+) colonies were selected as invasives and used for further characterization of the strains. Forty-one *S. flexneri* strains from households with two or more individuals infected with this species were selected to further characterize and compare the strains as follows:

1. biochemical analysis using the API 20E system;
2. serological analysis where the strains identified as shigella were sent to LANARE, Lima, Peru, for their corresponding serotyping;
3. antibiotic susceptibility testing which was done on Muller–Hinton agar plates by the disk diffusion method of Bauer et al. [26], antimicrobial agents tested included: ampicillin (A), cephalothinin (Cp), chloramphenicol (C), doxycycline (D), erythromycin (E), gentamicin (G), kanamycin (K), nalidixic acid (NA), neomycin (N), streptomycin (S), tetracycline (T), trimethoprim–sulphamethoxazole (Tsx), and triple sulfia (Ts);
4. analysis of virulence using the Sereny test [27], where approximately 10⁸ cells were inoculated into the conjunctival sac of a guinea-pig, and strains that elicited keratoconjunctivitis within 48 h were considered Sereny positive;
5. plasmid profile analysis performed by the method of Birnboim and Doly [28], where plasmids were detected by electrophoresis in 0.8% agarose gels containing 0.5 µg of ethidium bromide/ml and photographed with UV light illumination using a polaroid camera (Fotodyne, Inc., New Berlin, WI, USA);
6. *ipaH* hybridization patterns using non-radioactive and radioactive *ipaH* probes.

Plasmid DNA was prepared from 2 ml cultures using the method of Birnboim and Doly [28] or from 100–200 ml cultures by a modified version of the method described by Cassie and Denarie [14]. Approximating 1 µg of purified DNA plasmid was digested with *HindIII* or *SalI* (Boehringer Mannheim, Indianapolis, IN, USA) according to the manufacturer’s instructions. DNA was electrophoresed on 0.8% agarose gels containing 0.5 µg of ethidium bromide/ml and photographed with UV light illumination using a polaroid camera (Fotodyne, Inc.). The separated DNA fragments were transferred to nitrocellulose membranes (Schleicher & Schuell, Keene, NH, USA) by the method of Southern [28], and hybridized with non-radio labelled or radiolabelled synthetic *ipaH* oligoprobes. The non-radioactive *ipaH* oligoprobe used in this study has been previously described [16], with the genetic sequence 5′-d CtG GAG GAC ATT GCC CGG G-3′, where ‘t’ indicated a modification with suberyl-alkaline phosphatase (Molecular Biosystems, San Diego, CA, USA). Prehybridization was carried out in 5 × SSPE [5 × SSPE is 0.9 M NaCl, 50 mM NaPO₄ (pH 7.7), and 5 mM EDTA] with 1% SDS/0.5% polyvinylpyrrolidone (Sigma-Aldrich, St. Louis, MO, USA) containing 100 µg of yeast tRNA (Gibco-BRL, Life Technologies, Gaithersburg, MD, USA) per ml for 15 min at 42 °C; this was followed by hybridization of each filter in 3 ml of the same buffer for 30 min at 42 °C by using 5 µl of the *ipaH* oligoprobe (250 nm/ml). The filter was washed twice at 50 °C for 10 min each time in 1 × SSPE/1% SDS for the first wash and in 1 × SSPE/1% Triton X-100 for the second wash. The third wash in 1 × SSPE was carried out at room temperature for 10 min. Detection of bound probe was accomplished by incubation of filters for 15 h.
at room temperature in the dark in 5 ml of 0·1 M Tris–HCl (pH 9·5)/0·1 M NaCl/50 mM MgCl₂ as the alkaline phosphatase substrate with 0·17 mg of 5-bromo-4-chloro-3-indolyl phosphate p-toluidine salt (BCIP) and 0·33 mg of nitroblue tetrazolium chloride (NBT) per ml. Substrate development was terminated by washing the filters in distilled water, then the filters were allowed to dry in the dark. The same ipaH oligoprobe was end labelled with [γ-32P]dATP by using T4 polynucleotide kinase (5'-DNA terminus labelling kit, Gibco-BRL) as previously described [10, 13]. The filters were hybridized and further processed as outlined previously [13]. Briefly, hybridizations were carried out for 16 h at 37 °C in 1 × SSPE, 0·2% SDS, 5 × Denhardt’s solution, 0·01% NaPPi and 100 μl/ml denatured calf-thymus DNA. After washing the filters several times with 1 × SSPE/0·2% SDS/0·01% NaPPi, the filters were exposed to X-ray film xAR (Eastman Kodak Co., Rochester, NY, USA).

Numerical analysis
The hybridization profiles of 41 S. flexneri strains were obtained after restriction enzyme digestion of their plasmid with HindIII and SalI. These profiles were used to generate a 0–1 character-state matrix for computer-based data storage and processing. One was scored when a band was present and 0 when absent. A dissimilarity matrix was derived from the character-state matrix by calculating Jaccard’s distance [29]. Classification was performed with a clustering algorithm according to the unweighted-pair group method with arithmetic averages (UPGMA) and dendrograms were produced by using a computer program elaborated by E. Gerres (Laboratoire d’Écologie Médicale et Pathologie Parasitaire, Faculté de Médecine Montpellier, France) [30].

Statistical analyses
Statistical analysis was done by Student’s t test using the INSTAT statistical analysis package (Graph Pad Software, Inc., San Diego, CA, USA). Significance was taken as P < 0·05.

RESULTS
Distribution of Shigella in Canto Grande
During the year of the study, 37 family studies were initiated and a total of 323 houses were visited in Canto Grande, an endemic area of shigellosis. In 109 houses (34%), at least one individual was detected with shigellosis, and in 34 of these houses (31%), between 2 and 4 individuals were infected with Shigella spp. A total of 161 shigella isolates were obtained and the distribution by species was as follows: 117 S. flexneri (73%), 21 S. boydii (13%), 15 S. dysenteriae (9%), and 8 S. sonnei (5%). Between 1 and 10 isolates (mean 4 isolates per family study) and between 1 and 6 serotypes (mean 3 serotypes per family study) were detected per family study, indicating a diversity of strains circulating in this population. One of 37 family studies showed the highest variation with 7 isolates of 6 different serotypes of shigella that were isolated. These 6 serotypes in this family study included S. flexneri 1a, S. flexneri 2a (2 isolates), S. flexneri 3a, S. flexneri 4a, S. flexneri 6 and S. boydii 11.

Comparison of S. flexneri in households with multiple infections
Since S. flexneri strains were isolated at a higher frequency in this study than other shigella serogroups (73%, P < 0·05), this species was selected to compare households with multiple shigella infections. The isolates of each household were compared by pairs. With two shigella isolates from a household, one intra-family comparison per pair (icp) was obtained. In a household with three shigella isolates, three combinations per pair were obtained (3 icp). In this way, 41 S. flexneri strains from 19 households with multiple shigella infections generated 25 icp cases. These icp cases were further characterized and divided into three groups depending on their similarities or differences in serotype, antibiogram susceptibility and plasmid profile. Group I consisted of 11 icp cases (44%) that showed similar characteristic for all markers used. There was only one exception in this group. Strains 30V101 and 30V103 belonging to S. flexneri serotype 1a were different in one out of 13 antibiotics tested (chloramphenicol, see Table 1). Group II contained 8 icp cases (32%), which were distinguished by different serotypes, plasmid profiles and antibiotic susceptibilities. One case included in this group showed the same antibiogram profile for the 13 antibiotics tested (3NC206, S. flexneri 6 and 3NC207, S. flexneri 2a, Table 1). Group III consisted of 6 icp cases (24%) with the same serotype but with different plasmid profile and antibiotic susceptibility (see Table 1). This indicated that (i) there were different shigella clones circulating at the same time in Canto...
Table 1. *Intra-family comparison by pairs (icp) of the S. flexneri strains from households with multiple infections*

<table>
<thead>
<tr>
<th>icp</th>
<th>Isolate</th>
<th>Sero-type</th>
<th><em>ipaH</em> pattern</th>
<th>API 20E</th>
<th>Antibiotic resistance*</th>
<th>Plasmid profile (MDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td><em>SalI</em></td>
<td><em>HindIII</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1NC101</td>
<td>5</td>
<td>H h</td>
<td>0044100</td>
<td>E, S, Ts, Tsx, S, Tsa</td>
<td>140, 4, 2-6, 2§</td>
</tr>
<tr>
<td>2</td>
<td>1NC103</td>
<td>5</td>
<td>H h</td>
<td>0044100</td>
<td>E, S, Ts, Tsa</td>
<td>140, 4, 2-6, 2</td>
</tr>
<tr>
<td></td>
<td>10VI00</td>
<td>6</td>
<td>J j</td>
<td>004102</td>
<td>E, T, A, S, Ts, Tsa</td>
<td>140, 4, 4, 2-6, 2</td>
</tr>
<tr>
<td></td>
<td>10VI05</td>
<td>6</td>
<td>J j</td>
<td>004102</td>
<td>E, D, T, A, S, Ts</td>
<td>140, 4-4, 4, 2-6, 2</td>
</tr>
<tr>
<td>3</td>
<td>10NC303</td>
<td>2a</td>
<td>C c</td>
<td>004100</td>
<td>D, T, C, S, Ts</td>
<td>140, 2-6, 2, 1-4</td>
</tr>
<tr>
<td></td>
<td>10NC306</td>
<td>2a</td>
<td>D d</td>
<td>004100</td>
<td>D, T, C, S, Ts</td>
<td>140, 2-6, 2, 1-4</td>
</tr>
<tr>
<td>4</td>
<td>15VI100</td>
<td>3a</td>
<td>G g</td>
<td>0004140</td>
<td>†</td>
<td>140, 3-4, 2-6, 2</td>
</tr>
<tr>
<td></td>
<td>15VI04</td>
<td>3a</td>
<td>G g</td>
<td>0004140</td>
<td></td>
<td>140, 3-4, 2-6, 2</td>
</tr>
<tr>
<td>5</td>
<td>16NC207</td>
<td>6</td>
<td>J j</td>
<td>004102</td>
<td>E, D, T, C, S, Ts</td>
<td>140, 4-4, 4, 1-8</td>
</tr>
<tr>
<td></td>
<td>16NC208</td>
<td>6</td>
<td>J j</td>
<td>004102</td>
<td>E, D, T, C, S, Ts</td>
<td>140, 4-4, 4, 1-8</td>
</tr>
<tr>
<td>6</td>
<td>20NC307</td>
<td>1a</td>
<td>A a</td>
<td>004100</td>
<td>E, T, C, A, S, Ts</td>
<td>140, 2-6, 2</td>
</tr>
<tr>
<td></td>
<td>20NC309</td>
<td>1a</td>
<td>A a</td>
<td>004100</td>
<td>E, T, C, A, S, Ts</td>
<td>140, 2-6, 2</td>
</tr>
<tr>
<td>7</td>
<td>23CI307</td>
<td>3a</td>
<td>G g</td>
<td>0044100</td>
<td>E, A, Tsx, S, Ts</td>
<td>140, 3-4, 2-6, 2, 1-4</td>
</tr>
<tr>
<td></td>
<td>23CI308</td>
<td>3a</td>
<td>G g</td>
<td>0044100</td>
<td>E, A, Tsx, S, Ts</td>
<td>140, 3-4, 2-6, 2, 1-4</td>
</tr>
<tr>
<td>8</td>
<td>24NC312</td>
<td>2a</td>
<td>C c</td>
<td>004100</td>
<td>E, T, C, S, Ts</td>
<td>140, 2-6, 2</td>
</tr>
<tr>
<td></td>
<td>26NC305</td>
<td>1a</td>
<td>A a</td>
<td>004100</td>
<td>T, C, A, S, Ts</td>
<td>140, 2-6, 2</td>
</tr>
<tr>
<td>9</td>
<td>30VI10</td>
<td>1a</td>
<td>A a</td>
<td>004100</td>
<td>T, A, S, Ts</td>
<td>140, 2-6, 2</td>
</tr>
<tr>
<td></td>
<td>30VI03</td>
<td>1a</td>
<td>A a</td>
<td>004100</td>
<td>T, C, A, S, Ts‡</td>
<td>140, 2-6, 2</td>
</tr>
<tr>
<td>10</td>
<td>31CI103</td>
<td>2a</td>
<td>C c</td>
<td>000412</td>
<td>D, T, C, A, S</td>
<td>140, 2-6, 2</td>
</tr>
<tr>
<td></td>
<td>31CI105</td>
<td>2a</td>
<td>C c</td>
<td>000412</td>
<td>D, T, C, A, S</td>
<td>140, 2-6, 2</td>
</tr>
<tr>
<td>11</td>
<td>3NC206</td>
<td>6</td>
<td>J j</td>
<td>004102</td>
<td>D, T, C, S, Ts</td>
<td>140, 2-6, 2</td>
</tr>
<tr>
<td></td>
<td>3NC207</td>
<td>2a</td>
<td>C c</td>
<td>004102</td>
<td>D, T, C, S, Ts</td>
<td>140, 2-6, 2</td>
</tr>
<tr>
<td></td>
<td>3NC302</td>
<td>2a</td>
<td>C c</td>
<td>004102</td>
<td>D, T, C, S, Ts</td>
<td>140, 2-6, 2</td>
</tr>
<tr>
<td></td>
<td>3NC305</td>
<td>6</td>
<td>J j</td>
<td>004102</td>
<td>E, D, C, K, S, N</td>
<td>140, 4-2, 1-8</td>
</tr>
<tr>
<td>12</td>
<td>5CI105</td>
<td>3a</td>
<td>G g</td>
<td>004100</td>
<td>E, D, T, C, A, K, Tsx, S, N, Ts</td>
<td>140, 3-4, 2-6, 2</td>
</tr>
<tr>
<td></td>
<td>5CI106</td>
<td>2a</td>
<td>C c</td>
<td>004100</td>
<td>E, D, T, C, A, K, S, N, Ts</td>
<td>140, 4-8, 3-4, 2-6, 2</td>
</tr>
<tr>
<td>13</td>
<td>6NC104</td>
<td>6</td>
<td>J l</td>
<td>004100</td>
<td>A</td>
<td>140, 4-2, 1-8, 1-4</td>
</tr>
<tr>
<td></td>
<td>6NC108</td>
<td>3a</td>
<td>G g</td>
<td>004100</td>
<td>E, D, T, C, A, S, Ts</td>
<td>140, 3-4, 2-6, 2</td>
</tr>
<tr>
<td>14</td>
<td>20NC305</td>
<td>3a</td>
<td>G g</td>
<td>004100</td>
<td>E, D, T, C, A, S, Ts</td>
<td>140, 3-4, 2-6, 2</td>
</tr>
<tr>
<td></td>
<td>20NC307</td>
<td>1a</td>
<td>A a</td>
<td>004100</td>
<td>E, T, C, A, S, Ts</td>
<td>140, 2-6, 2</td>
</tr>
<tr>
<td>15</td>
<td>20NC305</td>
<td>3a</td>
<td>G g</td>
<td>004100</td>
<td>E, D, T, C, A, S, Ts</td>
<td>140, 3-4, 2-6, 2</td>
</tr>
<tr>
<td></td>
<td>20NC309</td>
<td>1a</td>
<td>A a</td>
<td>004100</td>
<td>E, T, C, A, S, Ts</td>
<td>140, 2-6, 2</td>
</tr>
<tr>
<td>16</td>
<td>25VI10</td>
<td>6</td>
<td>K k</td>
<td>0004002</td>
<td>E, T, C, S, Ts</td>
<td>140, 4-2, 1-8</td>
</tr>
<tr>
<td></td>
<td>25VI08</td>
<td>1a</td>
<td>A a</td>
<td>004100</td>
<td>D, T, C, S, Ts</td>
<td>140, 2-6, 2</td>
</tr>
<tr>
<td>17</td>
<td>34VI100</td>
<td>1a</td>
<td>A a</td>
<td>004100</td>
<td>E, D, T, C, A, K, Tsx, S, N, Ts</td>
<td>140, 4-8, 3-4, 2-6, 2, 1-4</td>
</tr>
<tr>
<td></td>
<td>34VI06</td>
<td>2a</td>
<td>E e</td>
<td>004100</td>
<td>D, T, C, S, Ts</td>
<td>140, 2-6, 2, 1-4</td>
</tr>
<tr>
<td>18</td>
<td>24CI204</td>
<td>1a</td>
<td>B b</td>
<td>004100</td>
<td>E, D, T, C, S, Ts</td>
<td>140, 90, 2-6, 2</td>
</tr>
<tr>
<td></td>
<td>24CI208</td>
<td>1a</td>
<td>B b</td>
<td>004100</td>
<td>E, D, T, C, S, Ts</td>
<td>140, 90, 2-6, 2</td>
</tr>
</tbody>
</table>

* Antibiotics: A, ampicillin; C, chloramphenicol; Cp, cephalothin; D, doxycycline; E, erythromycin; G, gentamicin; K, kanamycin; NA, nalidixic acid; N, neomycin; S, streptomycin; T, tetracycline; Tsx, trimethoprim–sulphamethoxazole; Ts, triple sulf.

† These two strains were sensitive to the 13 antibiotics tested.
‡ Due to the difference in only one antibiotic, these strains were included in Group I.
§ Sizes of the plasmids refer to the sizes of the supercoiled form of the plasmid.
Grande, and (ii) *S. flexneri* strains were identical in less than half (44%) of the households with multiple shigella infections.

**ipaH hybridization profiles of *S. flexneri* strains from households with multiple shigella infections**

The 25 icp cases distributed in three groups described above were compared utilizing their *ipaH* hybridization patterns. The laboratory strain M90T (*S. flexneri* 5) was included as a control for the hybridizations with the *ipaH* oligoprobe. The hybridization patterns were similar using either non-radioactive (Fig. 1a) or radioactive (Fig. 1b–d) *ipaH* probes and these patterns are schematized in Figure 2(a, b) (*SalI* and *HindIII* restriction enzyme-digested DNA respectively). The *ipaH* patterns of the three *S. flexneri* groups described previously were analysed. We arbitrarily used capital letters for *SalI* patterns and lowercase letters for *HindIII* patterns.

**Group I:** In 10 out of 11 icp cases (91%) the *ipaH* patterns were identical (Table 1). Only 1 icp case (icp no. 3, 10NC303 and 10NC306 corresponding to serotype 2a) showed a difference in 1 out of 5 bands...
Fig. 2. Schematic representation of the ipaH hybridization profiles in invasion plasmid DNA isolated from S. flexneri strains from households with multiple shigella infections. Invasion plasmid DNA was isolated as described in the Materials and Methods section and digested with SalI (a) or HindIII (b) prior to electrophoretic separation on 0.8% agarose gels and blotting to nitrocellulose. Samples were screened with the ipaH oligoprobe as described. Molecular-weight markers are indicated to the left of each panel and reflect the mobilities of λHindIII and ΦX174 HaeIII fragments.
when the DNA was digested either with *Sal*I (Fig. 1c, lanes 2 and 3, these patterns are also illustrated in Fig. 2a, serotype 2 patterns C and D) or *Hind*III (Fig. 1d, lanes 3 and 4, also illustrated in Fig. 2b, patterns c and d).

*Group II*: These cases were variable in their serotypes, antibiotic susceptibility and plasmid profiles. These strains also showed differences in their *ipaH* hybridization patterns (Table 1, Fig. 2a, b).

*Group III*: There were 6 icp cases in this group that had the same serotype but with differences in antibiotic susceptibility and plasmid profiles. In 5 out of 6 icp cases (83%) the *ipaH* pattern was the same when either *Hind*III or *Sal*I was used to digest the invasion plasmid (Table 1). Only 1 of the 6 icp cases showed a different *ipaH* hybridization profile. The *ipaH* profile of this icp case (10NC303 and 10NC308, *S. flexneri* 2a) has been described above for Group I.

A strong correlation was obtained particularly between the serotype of *S. flexneri* and *ipaH* patterns. These patterns are illustrated schematically in Figure 2(a, b). The distribution of the *ipaH* patterns, antibiotic susceptibility patterns and plasmid profiles of the 41 *S. flexneri* strains is further described in Table 2. With strains of serotype 1 (Fig. 2a, b), pattern A (*Sal*I) and pattern a (*Hind*III) were observed in 8 out of 10 strains, while pattern B (*Sal*I) and pattern b (*Hind*III) were observed in 2 strains from the same house (24CI204 and 24CI208) belonging to Group III (Table 1). Strains of serotype 2a which were more numerous (13 strains) exhibited greater variation in their *ipaH* hybridization patterns. Using *Sal*I, 4 patterns (C–F) were observed (Fig. 2a) while with *Hind*III-digested DNA, three patterns were obtained (Fig. 2b). The majority of strains of serotype 2a (8/13) presented patterns C and c. Two strains of this serotype which gave patterns D and d, were isolated from the same house and exhibited differences in plasmid profiles and antibiotic susceptibility (Tables 1 and 2). Only one strain of serotype 2a in this study presented the patterns E and e (Fig. 2a, b). Pattern F (Fig. 2a) and pattern e (Fig. 2b) were observed in two strains of serotype 2a and these two strains showed differences in antibiotic susceptibility and plasmid profiles (Tables 1 and 2). In the case of serotype 3a, the eight strains presented the same pattern (G and g; Fig. 2a, b). Interestingly, three strains of this serotype (23CI307, 23CI308 and 23CI309) were isolated from the same house, of which two (23CI307 and 23CI308) were identical but one (23CI309) showed differences

### Table 2. Comparison of *ipaH* patterns, antibiotic-resistance and plasmid profiles of *S. flexneri* serotypes

<table>
<thead>
<tr>
<th>Serotype</th>
<th><em>ipaH</em> pattern</th>
<th>No. of cases</th>
<th>Antibiotic resistance*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>8</td>
<td>ampicillin, chloramphenicol, cefuroxime, doxycycline, erythromycin, gentamicin, kanamycin, neomycin, streptomycin, tetracycline, trimethoprim–sulphamethoxazole, triple sulfa.</td>
</tr>
<tr>
<td>2a</td>
<td>C</td>
<td>4</td>
<td>ampicillin, chloramphenicol, cefuroxime, doxycycline, erythromycin, gentamicin, kanamycin, neomycin, streptomycin, tetracycline, trimethoprim–sulphamethoxazole, triple sulfa.</td>
</tr>
<tr>
<td>2a</td>
<td>C</td>
<td>3</td>
<td>ampicillin, chloramphenicol, cefuroxime, doxycycline, erythromycin, gentamicin, kanamycin, neomycin, streptomycin, tetracycline, trimethoprim–sulphamethoxazole, triple sulfa.</td>
</tr>
<tr>
<td>2a</td>
<td>C</td>
<td>1</td>
<td>ampicillin, chloramphenicol, cefuroxime, doxycycline, erythromycin, gentamicin, kanamycin, neomycin, streptomycin, tetracycline, trimethoprim–sulphamethoxazole, triple sulfa.</td>
</tr>
<tr>
<td>2a</td>
<td>D</td>
<td>2</td>
<td>ampicillin, chloramphenicol, cefuroxime, doxycycline, erythromycin, gentamicin, kanamycin, neomycin, streptomycin, tetracycline, trimethoprim–sulphamethoxazole, triple sulfa.</td>
</tr>
<tr>
<td>2a</td>
<td>C</td>
<td>2</td>
<td>ampicillin, chloramphenicol, cefuroxime, doxycycline, erythromycin, gentamicin, kanamycin, neomycin, streptomycin, tetracycline, trimethoprim–sulphamethoxazole, triple sulfa.</td>
</tr>
<tr>
<td>2a</td>
<td>C</td>
<td>1</td>
<td>ampicillin, chloramphenicol, cefuroxime, doxycycline, erythromycin, gentamicin, kanamycin, neomycin, streptomycin, tetracycline, trimethoprim–sulphamethoxazole, triple sulfa.</td>
</tr>
<tr>
<td>2a</td>
<td>F</td>
<td>2</td>
<td>ampicillin, chloramphenicol, cefuroxime, doxycycline, erythromycin, gentamicin, kanamycin, neomycin, streptomycin, tetracycline, trimethoprim–sulphamethoxazole, triple sulfa.</td>
</tr>
<tr>
<td>3a</td>
<td>G</td>
<td>4</td>
<td>ampicillin, chloramphenicol, cefuroxime, doxycycline, erythromycin, gentamicin, kanamycin, neomycin, streptomycin, tetracycline, trimethoprim–sulphamethoxazole, triple sulfa.</td>
</tr>
<tr>
<td>3a</td>
<td>G</td>
<td>2</td>
<td>ampicillin, chloramphenicol, cefuroxime, doxycycline, erythromycin, gentamicin, kanamycin, neomycin, streptomycin, tetracycline, trimethoprim–sulphamethoxazole, triple sulfa.</td>
</tr>
<tr>
<td>3a</td>
<td>G</td>
<td>2</td>
<td>ampicillin, chloramphenicol, cefuroxime, doxycycline, erythromycin, gentamicin, kanamycin, neomycin, streptomycin, tetracycline, trimethoprim–sulphamethoxazole, triple sulfa.</td>
</tr>
<tr>
<td>3a</td>
<td>H</td>
<td>2</td>
<td>ampicillin, chloramphenicol, cefuroxime, doxycycline, erythromycin, gentamicin, kanamycin, neomycin, streptomycin, tetracycline, trimethoprim–sulphamethoxazole, triple sulfa.</td>
</tr>
<tr>
<td>3a</td>
<td>J</td>
<td>6</td>
<td>ampicillin, chloramphenicol, cefuroxime, doxycycline, erythromycin, gentamicin, kanamycin, neomycin, streptomycin, tetracycline, trimethoprim–sulphamethoxazole, triple sulfa.</td>
</tr>
</tbody>
</table>

*Antibiotics: A, ampicillin; C, chloramphenicol; Cp, cephalothin; D, doxycycline; E, erythromycin; G, gentamicin; K, kanamycin; NA, nalidixic acid; N, neomycin; S, streptomycin; T, tetracycline; Tsx, trimethoprim–sulphamethoxazole; Ts, triple sulfa.
in plasmid profile, antibiotic-susceptibility pattern and biochemical reactions. Only two strains of serotype 5 were obtained in this study with identical markers. When the \(\text{ipaH}\) patterns of these \(S.\ flexneri\) 5 strains were compared to the pattern obtained with a laboratory strain control (M90T-W), differences in only one band were obtained with either \(SalI\) or \(HindIII\) digestions (Fig. 2a, b). In the case of serotype 6, eight strains were studied. Two patterns where obtained with \(SalI\) (J and K) and three patterns with \(HindIII\) (j, k, l) digestions (Fig. 2a, b). The pattern J (\(SalI\)) was observed in seven strains of serotype 6 (Table 2); however, in one of these strains the pattern ‘l’ instead of the pattern ‘j’ was obtained when \(HindIII\) was used for DNA digestion. Only one strain of serotype 6 showed pattern K (\(SalI\); Fig. 2a) and pattern k (\(HindIII\); Fig. 2b).

**Phylogenetic analysis**

Based on \(\text{ipaH}\) patterns of hybridization and serotype a dendrogram representing cluster analysis was carried out as described in Figure 3. The \(S.\ flexneri\) strains used in this study clustered in four groups. The first cluster includes the patterns Aa, Gg, Kk, Jj and Ji, of serotypes 1a, 3a and 6. The second cluster includes only two isolates of \(S.\ flexneri\) serotype 1a, 24C1204 and 24C1208, which presented the pattern Bb. These two isolates of \(S.\ flexneri\) 1a were biochemically different compared to the other eight strains of this serotype. The third cluster corresponds to the patterns Cc, Dd, Ee and Fc of the 13 \(S.\ flexneri\) 2a isolates. Finally, the fourth cluster includes the two isolates of \(S.\ flexneri\) serotype 5 with pattern Hh and the strain control M90T-W.

For the Group I strains described in Table 1 in this study, 10 out of 11 icp cases showed a Jaccard’s distance of 0 when the \(\text{ipaH}\) hybridization patterns were compared using numerical analysis. Only 1 icp case [strains 10NC303 and 10NC306 (\(S.\ flexneri\) 2a)] presented small differences in their respective \(\text{ipaH}\) patterns and had a Jaccard’s distance of only 0-1 (Fig. 3, Table 1). All of the icp cases in Group II, with strains that presented different serotypes and \(\text{ipaH}\) hybridization profiles, had Jaccard’s distances that varied between 0-31 and 0-5 (Fig. 3, Table 1). In Group III, 5 out of 6 icp cases presented a Jaccard’s distance of 0. Curiously, the strains (10NC303 and 10NC308, \(S.\ flexneri\) 2a) of the only icp case with a different \(\text{ipaH}\) pattern in this group had a Jaccard’s distance of 0-1 and were isolated from the same house of the other icp case in Group I that showed differences in the \(\text{ipaH}\) profiles (Fig. 3, Table 1).

**DISCUSSION**

In this study we found that a diversity of shigella clones were circulating in the population of Canto Grande at the time of infection. Up to six different serotypes of \(Shigella\) were isolated from the household contacts of one index case. This complicated picture of the epidemiology of shigella has also been observed in other endemic areas of shigellosis, such as Bangladesh and Somalia [3]. A total of 41 \(S.\ flexneri\) strains distributed in 25 icp cases were divided in three groups when their respective serotypes, plasmid and anti-biogram profiles and \(\text{ipaH}\) hybridization profiles were compared.

Several methods have been used to study the epidemiology of shigella infections. Serology constitutes one of the most important analyses for species identification of shigella isolates and is based on the specificity of the O-antigen [31]. However, this technique shows limitations since it is time consuming, does not always provide reliable data, and needs extensive confirmation especially with non-typable strains [32]. Additionally, serological analysis cannot be considered alone since strains of the same serotype can be obtained from different geographical regions, which have been exposed to pressure-selection changes in both their phenotypic and genotypic properties [33]. Plasmid fingerprinting [33–40], antibiotic profiles [33, 34, 38, 40–42], analysis of the rRNA gene restriction patterns [43–47], phage typing [48, 49], genomic DNA analysis by restriction fragment length polymorphism (RFLPs) [20, 50], pulsed-field gel electrophoresis (PFGE) of DNA samples [32, 35, 45, 51] and distribution of insertion sequences [32] have provided alternative techniques for shigella epidemiology. Individually, each assay has limitations.

In this study we exploited the \(\text{ipaH}\) heterogeneity as a tool for distinguishing shigella strains in an epidemiological study of transmission in an endemic area of shigellosis. We found a strong correlation between \(\text{ipaH}\) patterns and serotypes of \(S.\ flexneri\). Our results indicate that \(\text{ipaH}\) hybridization patterns can differentiate between \(S.\ flexneri\) strains and can complement serotyping. \(\text{ipaH}\) genes are detected in multiple copies on the chromosome and invasion plasmids of virulent and avirulent \(Shigella\) spp. and EIEC strains [10]. The RFLP \(\text{ipaH}\) patterns not only correlated with the species and the geographic origin of the
Fig. 3. Dendrogram representing cluster analysis of the *ipaH* hybridization profiles and serotypes of the *S. flexneri* strains from households with multiple shigella infections.
strains, but were also able to discriminate between shigella isolates that had previously been shown to have similar RFLP patterns using Shiga toxin or sRNA probes [10]. In our study, we were able to discriminate between isolates of the same serotype. For example, we found that three isolates of S. flexneri 2a (strains 10NC303, 10NC306 and 10NC308) from the same household presented differences in their respective ipaH hybridization profiles. The three icp cases generated with these isolates were either identical or had some differences in their plasmid and antibiotic profiles. This indicates that in endemic areas such as Canto Grande, the shigella strains are undergoing rapid changes. Mutations or rearrangements in a region of the invasion plasmid will generate a different ipaH profile, even though the antibiotic characteristics and the presence of other plasmids are conserved (Group I, case no. 3). Alternatively, the ipaH profile can be conserved but changes in plasmids and/or antibiotic susceptibility may occur (Group III, case no. 21). We do not know which event occurred first or whether both events leading to differences among isolates occurred simultaneously. These changes result in the generation of an isolate with distinct phenotypic and genetic characteristics (Group III, case no. 20). A limitation of this study was to select and save only one strain per individual. Later, in another study of shigellosis conducted in Lima and Canto Grande, three shigella colonies per plate of selective medium were stored. We found that an individual could be infected with two different species of shigella, S. flexneri and S. sonnei (Fernandez-Prada CM, unpublished observations), and in some cases where there was only one species isolated, up to three clonal variants were detected when their ipaH RFLP profiles and serology were compared (Hartman AB, unpublished observations). These findings indicate that several clones are circulating in endemic areas of shigellosis and that the use of ipaH RFLP analysis of shigella strains is a powerful tool in epidemiological studies.

The role of ipaH in pathogenesis is unclear. Using suicide vector-driven mutagenesis, deletions in individual copies of ipaH8, ipaH13, and ipaH14 genes have been made, and double and triple ipaH mutants have been constructed [52]. These ipaH mutants remain invasive in HeLa cells and behave like wild-type strains in plaque assay indicating that these genes may not be involved in the entry or dissemination process [52]. However, these mutations induce an exaggerated Sereny response in guinea-pig eyes suggesting that ipaH may play a role in modulating the inflammatory response elicited by infection [52]. Infection of human monocyte-derived macrophages in vitro with virulent S. flexneri results in cell death with release of IL-1β [53, 54]. Recently, we have found that the rapid exit of virulent shigella from the phagocytic vacuoles of monocyte/macrophage cells may be facilitated by ipaH [52]. In the absence of IpaH, reduced cytotoxicity and lower levels of IL-1β secretion were observed [52].

Comparison of the plasmid profiles of the 41 S. flexneri strains studied here, indicated that the 2-6 and 2-0 MDa plasmids were present in 95 and 88% of the isolates respectively. As shown in Table 2, these plasmids were absent only in some strains of serotype 6. The presence of these small plasmids in S. flexneri strains has been previously reported and indicates that a correlation between plasmid profiles and serotypes for S. flexneri exists [36, 37, 39]. Additionally, we found that only the serotypes 5 and 6 of S. flexneri had the 4 MDa plasmid, while the 3-4 MDa plasmid was present in all 8 S. flexneri 3a, in 1 out of 13 S. flexneri 2a and in 1 out of 10 S. flexneri 1a strains. In S. flexneri strains of serotype 6, when the 2-0 MDa plasmid was not present a small plasmid of 1-8 MDa was observed instead. The S. flexneri strains analysed in this study were multi-drug resistant and contained several plasmids; however, we did not find a correlation between presence of a particular plasmid and antibiotic-resistance profile. Multi-drug resistance patterns were observed in isolates containing both large-sized plasmids and those containing only small-sized plasmids in addition to the 140 MDa invasion plasmid. Several reports have indicated the presence of large transmissible plasmids that harbour multiple resistance genes among isolates of shigella [33–36, 38, 40, 42]. Recently, Rajakumar et al. [55, 56] have identified a chromosomal region in S. flexneri strains that confers multi-antibiotic resistance. A 99-kb deletion within the chromosome of S. flexneri 2a (strain YSH6000) indicated that this region was responsible for susceptibility to ampicillin, streptomycin, tetracycline and chloramphenicol [56]. This 99-kb chromosomal multi-antibiotic resistance locus was highly similar in sequence and organization to a region of the Shigella R-plasmid, NR1, and it may have arose following integration of an NR1-like plasmid [55]. The possibility of integration of R-plasmids in the chromosome of shigella strains in Canto Grande is an open question and could be the explanation for the differences in antibiotic and plasmid profiles.
observed between isolates of icp cases with the same serotype and similar *ipaH* hybridization profiles (Group III). Curiously, the strains 15VI00 and 15VI04 (see Table 1, icp no. 4, Group I) were sensitive to the 13 antibiotics tested and this may be explained by a spontaneous deletion of the integrated multi-drug resistance genes in the chromosome of shigella or by the loss of the non-integrated R-plasmid.

Shigellosis remains endemic throughout the world, especially in many developing countries where poor sanitary conditions, malnutrition, overcrowded situations and unclean water supplies contribute to the high incidence and persistence of infection [2]. Shigellosis is a disease that can be a major killer of infants, children, elderly people, malnourished individuals and people living in marginal conditions [2] as in the case of Canto Grande in Lima, Peru. In order to understand the epidemiology of shigella, especially in endemic areas, it is necessary to have effective discriminatory tools to differentiate between isolates of shigella. These epidemiological tools have to be both easy to use and less time-consuming. We found that the *ipaH* hybridization profiles together with the data generated from serology, plasmid patterns and antibiotic characteristics represent a useful epidemiological tool not only to differentiate between strains of *S. flexneri* but also to understand the dynamic of the generation of new clones of pathogenic bacteria.

ACKNOWLEDGEMENTS

This study was supported by the Programme for the Control of Diarrhoeal Diseases of the World Health Organization, and by the Applied Diarrhoeal Disease Research Project of the United States Agency for International Development. The authors thank the field and laboratory workers and supervisors of the study for their dedicated assistance, especially Dora Maurtua, Rina Meza, Nancy Quevedo and Cesar Guerrero. Thanks are also due to the Laboratorio Nacional de Referencia de Enteropatogenos (LANARE) in Lima, Peru for the serotyping of the shigella strains, and to Dr David Taylor for reading the manuscript.

REFERENCES


Downloaded from https://www.cambridge.org/core. IP address: 54.70.40.11, on 24 Jan 2019 at 23:56:35, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. 
https://doi.org/10.1017/S0950268803001560
Molecular epidemiology of shigellosis in Peru


