Abbreviations: EIEC, enteroinvasive Escherichia coli; MLEE, multilocus enzyme electrophoresis; RAPD, random amplified polymorphic DNA; RFLP, restriction fragment length polymorphism.

INTRODUCTION

DNA hybridization studies have shown that Shigella (with the exception of Shigella boydii serotype 13) and Escherichia coli belong to the same genetic species (Brenner et al., 1972, 1973, 1982). However, nutritional characterization, numerical taxonomy and enzyme electrophoresis show that there are differences between these organisms (Johnson et al., 1975; Veron & Le Minor, 1975; Ochman et al., 1983) and within Shigella, between the four taxospecies, Shigella dysenteriae, Shigella flexneri, S. boydii and Shigella sonnei (Goullet 1980; Dodd & Jones, 1982). A previous analysis of esterase electrophoretic polymorphisms distinguished five clusters among the Shigella: (1) S. dysenteriae serotype 1; (2) S. flexneri serotypes 1–5; (3) S. flexneri serotype 6 and S. boydii serotypes 2 and 4; (4) S. sonnei; and (5) S. boydii serotype 13. These clusters are more closely related to E. coli than to each other (Goullet & Picard, 1987).

The concept of a clonal structure was developed by Selander & Levin (1980) for E. coli and subsequent studies have shown that the linkage disequilibrium in this species is very close to the theoretical maximum, indicating that recombination events are rare. Moreover, the clonal structure of the species was demonstrated by the similarity obtained within the ECOR collection of 72 E. coli strains for phylogenetic trees based on three kinds of data (Ochman & Selander, 1984; Desjardins et al., 1995). The four phylogenetic

Shigella and enteroinvasive Escherichia coli strains are derived from distinct ancestral strains of E. coli

Karine Rolland,1 Nicole Lambert-Zechovsky,2 Bertrand Picard1 and Erick Denamur3

Author for correspondence: Bertrand Picard. Tel: +33 2 98 01 64 56. Fax: +33 2 98 01 64 74. Bertrand.Picard@univ-Brest.fr

The differentiation between Shigella subspecies, and the phylogenetic position of Shigella clones within Escherichia coli clones was determined by analysis of restriction fragment length polymorphisms of rDNA (ribotyping). Seventy-five Shigella strains belonging to the four subspecies and 13 enteroinvasive E. coli (EIEC) strains were compared with the 72 E. coli strains of the ECOR collection, which have been classified into four phylogenetic groups (A, B1, B2 and D). Seventeen Shigella dysenteriae ribotypes, 12 Shigella flexneri ribotypes, 23 Shigella boydii ribotypes, 12 Shigella sonnei ribotypes and 13 EIEC ribotypes were identified following digestion with HindIII and EcoRI. Correspondence analysis of the data showed that S. boydii serotype 13 strains were distantly related to the other Shigella strains, and that S. sonnei and S. flexneri were distinct from S. boydii and S. dysenteriae. The ribotypes of Shigella and ECOR strains were indistinguishable, and S. sonnei, S. flexneri and most S. dysenteriae strains were closely related to phylogenetic group D, whereas S. dysenteriae serotype 1 strains belonged to phylogenetic group B1, and S. boydii strains were evenly distributed between the two groups. The Shigella strains were distantly related to group B2, which contains E. coli strains frequently implicated in extra-intestinal infections in humans. In contrast, the 13 EIEC strains were more widely distributed between phylogenetic groups B1, A and B2. Thus, there was no primordial Shigella species and Shigella and EIEC strains are derived from different ancestral strains.

Keywords: Shigella, Escherichia coli, enteroinvasive Escherichia coli (EIEC), phylogeny
groups A, B1, B2 and D, described by Herzer et al. (1990) on the basis of multilocus enzyme electrophoresis (MLEE), are also distinguishable by random amplified polymorphic DNA (RAPD), and restriction fragment length polymorphism (RFLP) analysis of rDNA genes (ribotyping). Thus ribotyping can be useful for classification of the E. coli and Shigella clones.

We first investigated the differentiation between the Shigella clusters by rDNA gene RFLP analysis of Shigella strains belonging to the four taxospecies. We then determined the phylogenetic position of the Shigella clones and of a group of enteroinvasive E. coli (EIEC) strains within the ECOR collection. This allowed us to assign the Shigella and EIEC strains to the previously described E. coli phylogenetic groups A, B1, B2 and D (Herzer et al., 1990).

METHODS

Bacterial strains

Shigella strains. The strains tested are listed in Table 1. We used 22 S. dysenteriae strains, 13 S. flexneri strains, 24 S. boydii strains (including two serotype 13 strains) and 16 S. sonnei strains.

E. coli strains. The 72 ECOR strains were obtained from R. K. Selandier, Pennsylvania State University, USA and the 13 EIEC strains were obtained from Dr P. Bouvet, Institut Pasteur, France (strains CNRSS 84-293, 85-90, 85-200, 85-206, 85-209, 85-242, 85-255, 94-8778, 95-7511 and 95-11794) and Dr Sansonetti, Institut Pasteur, France (strains 4608-0124, 1-72 and 13-80). The EIEC were shown to be enteroinvasive by the Sereny test (Sereny, 1955) for three strains or by hybridization using DNA probes (Bohnert et al., 1988) for the others. Among the EIEC, two were motile.

RFLP of rDNA regions. Bacterial DNA was prepared as previously described (Picard-Pasquier et al., 1989). DNA was digested with EcoRI and HindIII restriction endonucleases (Boehringer) according to the manufacturer’s instructions and was analysed by electrophoresis on submerged 0.8% agarose gels containing ethidium bromide. DNA fragment size markers (marker II; Boehringer) were used. Separated DNA restriction fragments were transferred to a nylon membrane (Amershams) by Southern blotting (Southern, 1975). Ribosomal 16S and 23S RNA from E. coli (Boehringer) was labelled by random oligopriming using a chemiluminescence system (Boehringer) as previously described (Bingen et al., 1992). Hybridization, washing and detection procedures were as previously described (Bingen et al., 1992).

Statistical analysis. The data were summarized in two two-way tables containing 72 rows for the 72 ECOR strains and 88 rows for the 75 Shigella strains and the 13 EIEC strains. The number of columns corresponded to the number of rDNA fragments produced by EcoRI and HindIII endonuclease digestion. For each column, the rDNA fragment was coded using a binary code, present = 2 or absent = 1, for each strain.

Correspondence analysis (Greenacre, 1984; Lebart et al., 1984; Tenehaus & Young, 1985) was performed for Shigella strains using one table. The other table was used for correspondence analysis of ECOR strains, in which the strains of Shigella and EIEC were treated as supplementary observations and were projected onto the factorial plane F1/F2 (see Fig. 3) obtained from the ECOR strains. Calculations were carried out using SPAD.N software.

RESULTS

RFLP analysis of rDNA regions

S. dysenteriae. rDNA was digested with EcoRI and 20 rDNA fragments 2–24 kb in size (Fig. 1a) were obtained from the 22 strains. We distinguished 17 rDNA RFLP patterns. When rDNA was digested with HindIII, fragments of 1.9–23.1 kb (Fig. 1b) were obtained and 10 rDNA RFLP patterns were identified. A combination of patterns resulting from both digestions was used to define ribotypes. Seventeen ribotypes were identified among the strains.

S. flexneri. EcoRI digestion produced 19 rDNA fragments 2–24 kb in size (Fig. 1a). There were 12 rDNA RFLP patterns in the 13 strains. HindIII digestion produced 18 rDNA fragments 1.9–23.1 kb in size which distinguished nine rDNA RFLP patterns (Fig. 1b). Twelve ribotypes were identified among the strains.

S. boydii. EcoRI digestion gave 23 rDNA fragments 2–24 kb in size (Fig. 1a), which distinguished 22 rDNA RFLP patterns among the 24 strains. HindIII digestion gave 21 rDNA fragments 1.9–23.1 kb in size (Fig. 1b), which distinguished 16 rDNA RFLP patterns. Twenty-three ribotypes were identified among the strains.

S. sonnei. EcoRI digestion produced 19 rDNA fragments with sizes of 2–24 kb (Fig. 1a), distinguishing 11 rDNA RFLP patterns. HindIII digestion produced 12 rDNA fragments 1.2–23.1 kb in size (Fig. 1b), which distinguished three rDNA RFLP patterns. Twelve ribotypes were identified among the strains.

EIEC. EcoRI digestion produced 16 rDNA fragments with sizes of 2–24 kb. Thirteen rDNA RFLP patterns were distinguished. HindIII digestion gave eight rDNA fragments of 1.9–23.1 kb in size, which distinguished nine rDNA RFLP patterns. Thirteen ribotypes were identified among the 13 strains.

Statistical analysis

Correspondence analysis was carried out with the ribotyping data obtained by the combination of the two endonucleases for the 75 Shigella strains. The plane F1/F2, which accounted for 25.5% of the total variance, showed that S. sonnei and S. flexneri strains were clearly distinguished from S. boydii (except serotype 13) and S. dysenteriae strains by the negative values of the first axis (Fig. 2). The second axis differentiated S. sonnei strains from S. flexneri strains, and S. boydii strains from S. dysenteriae strains. S. boydii serotype 13 strains were clearly distinguished from the other Shigella strains by the negative values of the first axis and the positive values of the second axis.

Correspondence analysis was carried out with the ribotyping data obtained by combination of the two endonucleases for the 72 ECOR strains and the plane
Table 1. Strains of Shigella examined

<table>
<thead>
<tr>
<th>Strains of Shigella examined</th>
<th>S. dysenteriae</th>
<th>S. flexneri</th>
<th>S. boydii</th>
<th>S. sonnei</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strain no.</td>
<td>Serotype</td>
<td>Strain no.</td>
<td>Serotype</td>
<td>Strain no.</td>
</tr>
<tr>
<td>-----------------</td>
<td>--------</td>
<td>-------------</td>
<td>----------</td>
<td>-------------</td>
</tr>
<tr>
<td>D1 8.76</td>
<td>3</td>
<td>D12 67.81</td>
<td>2</td>
<td>F1 10.80</td>
</tr>
<tr>
<td>D2 15.77</td>
<td>1</td>
<td>D13 85.99</td>
<td>2</td>
<td>F2 14.79</td>
</tr>
<tr>
<td>D3 53.1</td>
<td>2</td>
<td>D14 93.46</td>
<td>7</td>
<td>F3 18.70</td>
</tr>
<tr>
<td>D4 30.73</td>
<td>5</td>
<td>D15 93.45</td>
<td>4</td>
<td>F4 6.74</td>
</tr>
<tr>
<td>D5 2.64</td>
<td>2</td>
<td>D16 93.51</td>
<td>10</td>
<td>F5 66.79</td>
</tr>
<tr>
<td>D6 5.77</td>
<td>2</td>
<td>D17 52.27</td>
<td>1</td>
<td>F6 302.78</td>
</tr>
<tr>
<td>D7 31.73</td>
<td>7</td>
<td>D18 53.136</td>
<td>1</td>
<td>F7 3.77</td>
</tr>
<tr>
<td>D8 12.77</td>
<td>1</td>
<td>D19 54.95</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>D9 3.76</td>
<td>2</td>
<td>D20 55.90</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>D10 6.77</td>
<td>4</td>
<td>D21 56.33</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>D11 48.37</td>
<td>5</td>
<td>D22 62.17</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

S. jiexneri:

<table>
<thead>
<tr>
<th>Strain no.</th>
<th>Serotype</th>
<th>Strain no.</th>
<th>Serotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1 10.80</td>
<td>6</td>
<td>F8 82.48</td>
<td>T 2a</td>
</tr>
<tr>
<td>F2 14.79</td>
<td>6</td>
<td>F9 52.24</td>
<td>4</td>
</tr>
<tr>
<td>F3 18.70</td>
<td>2</td>
<td>F10 52.36</td>
<td>1a</td>
</tr>
<tr>
<td>F4 6.74</td>
<td>6</td>
<td>F11 52.39</td>
<td>2b</td>
</tr>
<tr>
<td>F5 66.79</td>
<td>4</td>
<td>F12 52.40</td>
<td>3</td>
</tr>
<tr>
<td>F6 302.78</td>
<td>6</td>
<td>F13 52.44</td>
<td>5</td>
</tr>
<tr>
<td>F7 3.77</td>
<td>6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

S. boydii:

<table>
<thead>
<tr>
<th>Strain no.</th>
<th>Serotype</th>
<th>Strain no.</th>
<th>Serotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1 52.23</td>
<td>1</td>
<td>B13 E.20</td>
<td>2</td>
</tr>
<tr>
<td>B2 42.80</td>
<td>1</td>
<td>B14 E.24</td>
<td>5</td>
</tr>
<tr>
<td>B3 41.80</td>
<td>11</td>
<td>B15 E.25</td>
<td>9</td>
</tr>
<tr>
<td>B4 12.80</td>
<td>10</td>
<td>B16 E.33</td>
<td>9</td>
</tr>
<tr>
<td>B5 19.80</td>
<td>5</td>
<td>B17 E.35</td>
<td>1</td>
</tr>
<tr>
<td>B6 54.73</td>
<td>1</td>
<td>B18 E.36</td>
<td>5</td>
</tr>
<tr>
<td>B7 56.73</td>
<td>5</td>
<td>B19 93.21</td>
<td>11</td>
</tr>
<tr>
<td>B8 12.77</td>
<td>10</td>
<td>B20 85.92</td>
<td>14</td>
</tr>
<tr>
<td>B9 2.77</td>
<td>5</td>
<td>B21 5.16</td>
<td>11</td>
</tr>
<tr>
<td>B10 16.76</td>
<td>2</td>
<td>B22 56.18</td>
<td>11</td>
</tr>
<tr>
<td>B11 57.43</td>
<td>9</td>
<td>B23 76.114</td>
<td>13</td>
</tr>
<tr>
<td>B12 5.99</td>
<td>10</td>
<td>B24 76.115</td>
<td>13</td>
</tr>
</tbody>
</table>

S. sonnei:

<table>
<thead>
<tr>
<th>Strain no.</th>
<th>Biotype</th>
<th>Strain no.</th>
<th>Biotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1 217.77</td>
<td>a</td>
<td>S9 13.77</td>
<td>a</td>
</tr>
<tr>
<td>S2 2.77</td>
<td>a</td>
<td>S10 1.78</td>
<td>g</td>
</tr>
<tr>
<td>S3 110.77</td>
<td>a</td>
<td>S11 82.49</td>
<td>g</td>
</tr>
<tr>
<td>S4 4.76</td>
<td>d</td>
<td>S12 51.1</td>
<td>g</td>
</tr>
<tr>
<td>S5 46.70</td>
<td>a</td>
<td>S13 52.55</td>
<td>d</td>
</tr>
<tr>
<td>S6 5.75</td>
<td>d</td>
<td>S14 55.56</td>
<td>d</td>
</tr>
<tr>
<td>S7 241.77</td>
<td>d</td>
<td>S15 66.4</td>
<td>a</td>
</tr>
<tr>
<td>S8 15.77</td>
<td>d</td>
<td>S16 67.63</td>
<td>a</td>
</tr>
</tbody>
</table>

*Biotypes were defined according to Szturm-Rubinstein (1964).

Fig. 1. Representative RFLP patterns of Shigella and ECOR rDNA regions obtained with (a) HindII and (b) EcoRI digestion enzymes. Lanes: 1, strain ECOR (E) 1; 2, strain E29; 3, strain E50; 4, strain E63; 5, Shigella strain D2; 6, strain D10; 7, strain F3; 8, strain F5; 9, strain B6; 10, strain B10; 11, strain S5; 12, strain S8. The strains are numbered as in Table 1. Positions of size markers are indicated on the right in kb.

F1/F2 was obtained (Fig. 3). The 75 Shigella strains and the 13 EIEC strains, treated as supplementary observations, were projected onto this plane, which accounted for 34.85% of the total variance. The majority of the ECOR strains were classified within the four phylogenetic groups A, B1, B2 and D previously distinguished by Herzer et al. (1990) using MLEE and by Desjardins et al. (1995) using ribotyping and RAPD analysis. The Shigella strains were projected into the same area as the ECOR strains. This shows that their ribotypes are very close to those of E. coli strains. However, no Shigella strains were projected within the area of phylogenetic groups A and B2, and all Shigella strains were projected within the area of phylogenetic groups B1 and D. Seven of the eight S. dysenteriae serotype 1 strains were
projected into the area of group B1. S. sonnei, S. flexneri
and the other S. dysenteriae strains were mostly pro-
jected into the area of phylogenetic group D. S. boydii
strains were more widely distributed within these two
ECOR areas. Eight of the 13 EIEC strains studied
belonged to group B1, 4 to group A and 1 to group B2.

**DISCUSSION**

As with previous studies based on biochemical and
serological characters (Dodd & Jones, 1982) and electro-
phoretic polymorphism of esterases (Goullet, 1980;
Goullet & Picard, 1987), clusters of *Shigella* were
distinguished by ribotyping (Figs 2, 3). Thus, S. sonnei
and S. flexneri were clearly distinguished from S. boydii
and S. dysenteriae, whereas the distinction between S.
boydii and S. dysenteriae was less clear (Fig. 2). The
strains of *S. boydii* serotype 13, which belong to a
distinct DNA hybridization group (Brenner et al.,
1982), were clearly distinguished from the other *Shigella*
strains. Our rDNA RFLP data did not demonstrate the
close taxonomic relationship between *S. flexneri* sero-
type 6 strains and *S. boydii* strains previously shown on
the basis of several characters (Timakov et al., 1972;
Dodd & Jones, 1982; Goullet et al., 1983).

Ribotyping was highly discriminatory in the *Shigella*
subspecies as 17 ribotypes were identified in *S.
dysenteriae*, 12 in *S. flexneri*, 23 in *S. boydii* and 12 in *S.
sonnei*. Serotyping (Orskov et al., 1977), biotyping
(Marranzano et al., 1985) and isoenzyme analysis
(Ochman et al., 1983) suggest that *S. sonnei* is genetically
homogeneous. Our study is consistent with previous
work showing that ribotyping can discriminate between
*S. sonnei* clones (Karaolis et al., 1994). RAPD analysis of

the 75 *Shigella* strains using the 186 and 3386 primers
(Desjardins et al., 1995) was less able than ribotyping to
discriminate the four taxospecies (data not shown).

Previous studies have shown sufficient similarity be-
tween *E. coli* and *Shigella* to place both in the same
species (Brenner et al., 1972, 1973; Goullet, 1980; Hartl
& Dykhuizen, 1984; Ochman et al., 1983; Whittam et
al., 1983). The correspondence analysis comparing
ECOR ribotypes and *Shigella* ribotypes supports this
(Fig. 3). However, *Shigella* clones were more closely
related to certain phylogenetic groups of the ECOR
collection as they were unambiguously projected within
the B1 and D groups even though some strains were

---

**Fig. 2.** Correspondence analysis of the 75 *Shigella* strains using
RFLP analysis of the rRNA genes. Projection of the strains,
numbered as in Table 1, onto the plane F1/F2. For clarity, when
several strains are projected onto the same point, only one is
indicated.

**Fig. 3.** Correspondence analysis of the 72 ECOR strains with the
75 *Shigella* and 13 EIEC strains treated as supplementary
observations, using RFLP analysis of rRNA gene data. (a) The 72
ECOR strains (●), the 13 EIEC strains (△) and the 75 *Shigella*
strains (○) are projected onto the plane F1/F2. The phylogenetic
groups of ECOR strains, A, B1, B2 and D (Herzer et
al., 1990), are indicated by dotted lines. (b) Enlargement of
the area of groups B1 and D. The *Shigella* strains are numbered as
in Table 1 e, ECOR.
projected within the B1 group area bordering the A group area. Moreover, S. sonnei, S. flexneri and most S. dysenteriae strains were closely related to group D, whereas S. dysenteriae serotype 1 strains were closely related to group B1 and S. boydii strains were widely distributed among the B1 and D groups. The Shigella strains were unrelated to group A. There have been several studies of the relationship between a limited sample of Shigella isolates and E. coli strains. These studies have produced conflicting results. MLFE analysis was used to classify most Shigella electrophoretic types to E. coli group II (corresponding to phylogenetic groups D and B1 of Herzer et al., 1990) (Whittam et al., 1983; Ochman et al., 1983; Ochman & Selander, 1984). Pupo et al. (1997) used the same technique to group most Shigella strains in a single cluster in phylogenetic group A, with S. flexneri serotype 2 and 4 strains in group D. However, the same authors showed that mdb sequencing groups all Shigella strains except S. sonnei within the E. coli B1 group strains. Such differences may be due to a higher rate of recombination in Shigella than in E. coli strains (Desjardins et al., 1995; Guttmann, 1997) or to convergent evolution resulting from selective pressure linked to their strict human intestinal biotype (Maynard Smith, 1996). Overall, the data support the notion put forward by Goullet & Picard (1987) in a study of esterase electrophoretic polymorphism that there was no single primordial Shigella species and Shigella strains are derived from different ancestral strains of E. coli.

The Shigella clones were found to be distantly related to the B2 phylogenetic group. This is consistent with their virulence in humans. Thus, Shigella cause bacillary dysentery whereas phylogenetic group B2 consists of highly pathogenic strains, frequently implicated in extra-intestinal infections and rarely isolated from the intestinal microflora (Goullet & Picard, 1986a; Bingen, 1986b). In contrast, despite having comparable virulence to Shigella and a homologous virulence plasmid in common (Ménard et al., 1996), the EIEC were found to belong to three phylogenetic groups of E. coli, consistent with the results of Pupo et al. (1997). Most strains were closely related to group B1, four strains belonged to group A and one to group B2. The proportion of EIEC belonging to the B2 group was similar to the proportion of E. coli strains belonging to the phylogenetic B2 group in the normal intestinal microflora of healthy individuals (Goullet & Picard, 1986b).

It has been suggested that Shigella plasmid virulence genes were acquired from another genus, based on the fact that the A + T content of the plasmid genes (70–73%) is higher than that of the chromosomal genes (50%) (Hale, 1991). Endonuclease digestion and Southern hybridization showed that the virulence plasmids of Shigella and EIEC are essentially homologous but have different endonuclease patterns (Sansonetti et al., 1983). This suggests that these may have been a common ancestor molecule which has been subject to evolution. Phylogenetic plasmid analysis should help to determine how the virulence plasmid was acquired by Shigella and EIEC strains.

ACKNOWLEDGEMENTS

We thank Philippe J. Sansonetti and Philippe Bouvet for providing the EIEC strains, and Jacques Elion for constant encouragement.

REFERENCES


K. Rolland and Others


Received 1 December 1997; revised 6 May 1998; accepted 18 May 1998.