The large plasmids of Shiga-toxin-producing 
Escherichia coli (STEC) are highly variable 
genetic elements

Werner Brunder, Herbert Schmidt, Matthias Frosch and Helge Karch

Author for correspondence: Werner Brunder. Tel: +49 931 201 3981. Fax: +49 931 201 3445.
e-mail: wbrunder@hygiene.uni-wuerzburg.de

Institut für Hygiene und Mikrobiologie der Universität Würzburg, 
Josef-Schneider-Str. 2, 
D-97080 Würzburg, 
Germany

Shiga-toxin-producing Escherichia coli (STEC) of different serotypes are known to harbour large plasmids. The aim of this study was to investigate, using the example of the plasmid-encoded serine protease EspP, whether these plasmids are a uniform genetic element present in STEC. Examination of 201 diarrhoeagenic E. coli strains using a newly developed espP-specific PCR showed that espP is specific for STEC and present in 57% of STEC belonging to 16 different serotypes. The espP genes of the 16 STEC serotypes varied to a certain extent, as shown by nucleotide sequence and restriction enzyme analyses, but the DNA regions adjacent to the espP gene were completely different. When two further STEC-plasmid markers, the catalase-peroxidase gene katP and the enterohaemorrhagic E. coli-haemolysin gene EHEC-hlyA were included, many combinations of the three markers were found, depending in part on the serotype. In addition, strains possessing none of the three markers still harboured large plasmids. In the most prevalent STEC serogroup, 0157, it was observed that the plasmid of sorbitol-fermenting STEC 0157:H7 lacks the espP and katP genes although both genes are present in the plasmid of the non-sorbitol-fermenting STEC 0157:H7. The EHEC-hlyA gene, however, is present in both. In conclusion, this study shows that the large plasmids of STEC are not uniform genetic elements but heterogeneous in both their gene composition and arrangement.

Keywords: Shiga-toxin-producing Escherichia coli, plasmids, variability, serine protease, espP

INTRODUCTION

Shiga-toxin-producing Escherichia coli (STEC) are a well-known cause of severe human diseases such as haemorrhagic colitis and haemolytic–uraemic syndrome (HUS). STEC harbour plasmids of various sizes, the most common being 75–100 kbp (Bopp et al., 1987; Ostroff et al., 1989; Fratamico et al., 1993; Paros et al., 1993).

Plasmids are also found in other pathogroups of diarrhoeagenic E. coli. They often carry genes responsible for the expression of their group-specific pathogenicity, e.g. the heat-labile and heat-stable toxins of enterotoxigenic E. coli (ETEC), the invasion factors of enteroinvasive E. coli (EIEC) or the aggregative adherence fimbriae mediating the typical phenotype of enteroaggregative E. coli (EAEC) (for a review see Nataro & Kaper, 1998). Plasmids or plasmid-encoded virulence factors were also used to define and differentiate the E. coli pathogroups. For example, possession of the so-called EAF (EPEC adherence factor) plasmid is used to characterize enteropathogenic E. coli (EPEC) (Nataro & Kaper, 1998). However, identification or classification of such plasmids has often been based on their size or on the presence of a small DNA fragment in the plasmid using hybridization or PCR. Examples are a PCR for the detection of EAEC plasmids (Schmidt et al., 1995a), probes specific for the EPEC EAF plasmid (Nataro et al., 1983; Girón et al., 1993) or the CVD419 probe for the large plasmid of STEC (Levine et al.,
1987). Apart from these examples, there is little information on the variability of the plasmids, or on their mobility or interchangeability between *E. coli* strains of different patho- or serotypes. As a rule, only the plasmid of one or two prototype strains from a particular group has been studied and characterized to any extent.

Recently, the nucleotide sequence of the large plasmid of an STEC O157:H7 strain (RIMD 0509952), derived from an outbreak in Japan in 1996, was completely determined (Makino et al., 1998). Our group investigated, at both the molecular and functional level, plasmid pO157 of STEC O157:H7 reference strain EDL933, isolated during one of the first reported outbreaks of *E. coli* O157:H7 in 1982 (O’Brien et al., 1983; Riley et al., 1983). We showed that this prototype plasmid carried the genetic information for the EHEC (enterohaemorrhagic *E. coli*) haemolysin (Schmidt et al., 1995b, 1996) which is responsible for the so-called enterohaemolytic phenotype. In addition, we found genes encoding a periplasmic catalase-peroxidase (*katP*, Brunder et al., 1996) and a complex type II secretion system (*ett*; Schmidt et al., 1997). Recently, we described an extracellular serine protease, termed EspP, encoded by pO157 (Brunder et al., 1997). We showed that EspP is able to cleave pepsin and coagulation factor V and have suggested therefore that EspP might be an accessory virulence factor exacerbating haemorrhagic colitis. Djafari et al. (1997) reported a similar protein (*PssA*) encoded by a large plasmid of a bovine STEC O26:H1 strain and showed that it is cytotoxic for Vero cells. Comparison of the two sequences revealed eight base differences within the ORF. One of these led to an amino acid substitution at amino acid position 1113 where *PssA* had a valine residue instead of the phenylalanine residue in EspP.

In the present study, we investigated the prevalence of the *espP* gene in a large number of clinical STEC isolates of a wide variety of serotypes, and in other diarrhoeagenic *E. coli*. Furthermore, we analysed the diversity of the *espP* gene within and between STEC serotypes and its linkage with other markers of O157. The results of this study may contribute to our understanding of the structure and diversity of STEC and their plasmids.

**METHODS**

**Bacterial isolates.** A total of 156 STEC isolates belonging to 30 serotypes were investigated (see Table 1). The strains were isolated between January 1995 and December 1996 from individual human stool samples collected in Germany. A few well-characterized STEC reference strains and isolates from bovine faeces, meat and raw milk were included in the study. Serotyping of *E. coli* was performed according to Bockemühl et al. (1992). D-Sorbitol fermentation was examined on D-sorbitol-MacConkey agar (Oxoid).

Of the 83 STEC isolates belonging to serogroup O157, 48 possessed the H7 flagellar antigen whereas 35 were H-.

**General recombinant DNA techniques.** Plasmid DNA was purified with Nucleobond AX cartridges (Macherey-Nagel) according to the instructions of the supplier. Purification of DNA from agarose gels was performed using a Prep-A-Gene kit (Bio-Rad). Restriction enzymes and T4 DNA ligase were purchased from Gibco-BRL and New England Biolabs. Plasmid pK18 (Pridmore, 1987) was used as the cloning vector. Restriction enzyme analysis, ligation and transformation were conducted according to standard procedures (Sambrook et al., 1989). Southern-hybridization experiments were performed using a digoxigenin labelling and detection kit (Boehringer Mannheim) according to the manufacturer’s instructions.

**Nucleotide sequencing.** Nucleotide sequencing was performed with an automatic sequencer (model 373A, Applied Biosystems) as described previously (Schmidt et al., 1995b). Both strands of the DNA were sequenced stepwise using from patients with diarrhoea or asymptomatic carriers. We also investigated 20 STEC strains belonging to serogroup O126 (11 isolates carried H11, nine isolates were H-), 12 of which harboured the stx1 gene and eight the stx2 gene. Half of the strains in this serogroup were from patients suffering from HUS. Of eight strains of serogroup O111, five possessed the stx1 gene, two the stx2 gene and one both genes. Eleven isolates belonging to serotype O103:H2 were also tested, 10 of which harboured the stx1 gene alone and one both the stx1 and stx2 genes. One to three strains of STEC serogroups rarely isolated from humans were included in the study.

Strains of other diarrhoeagenic *E. coli* pathogroups (EPEC, STEC, EIEC and EAEC) were obtained from our strain collection. *E. coli* K-12 laboratory strains DH5a and HB101 (Gibco-BRL) were used as hosts for cloned DNA. STEC O157:H7 prototype strain EDL933 (O’Brien et al., 1983; Riley et al., 1993) and its plasmid-cured derivative EDL933-cu (Tzipori et al., 1987) were used as control strains for PCR and hybridization experiments.

**PCR.** PCR for detection of *espP* sequences was performed using the GeneAmp 9600 PCR System (Perkin Elmer). Primers esp-A (5’-AAACAGCAGGACCTTTGAAGC-3’) and esp-B (5’-GGAGTCTGATCGATGAGAT-3’) were designed to amplify a 1830 bp fragment of the *espP* gene. Amplification was carried out in a total volume of 10 μl containing 5 μl bacterial cell suspension (about 10^8^ bacterial cells suspended in 0.85%, w/v, sodium chloride), 30 pmol each primer, 200 μM each dNTP, 5 μl 10× GeneAmp PCR buffer II, 3 μl 25 mM MgCl2 and 2 units AmpliTaq DNA polymerase (Perkin Elmer). After an initial denaturation step of 5 min at 94 °C the samples were subjected to 30 cycles, each consisting of 30 s at 94 °C (denaturing), 60 s at 56 °C (annealing) and 150 s at 72 °C (extension). The reaction was completed with a final extension step of 5 min at 72 °C. The PCR products were analysed by agarose gel electrophoresis and ethidium bromide staining. Each reaction was conducted at least twice in independent reactions.

PCR for detection of the *katP* gene and the EHEC-hlyA gene was performed as described previously (Schmidt et al., 1995b; Brunder et al., 1996). PCRs for the investigation of DNA regions adjacent to the *espP* gene were performed using primers web-5 (5’-CAGAACTATATGCTTGGATGTA-3’) and web-6 (5’-AATATGCCTTTGCTTGTCG-3’), web-7 (5’-GGGAAACGGTGTGATGATG-3’), web-8 (5’-TCGATTTCGCTGCCTGACTCTG-3’) and web-9 (5’-TCGTTACGCAGCGGACCTTA-3’). See Fig. 3 for the binding sites of these primers. Reaction conditions for primer pairs web-5/web-6 and web-9/web-6 were identical to that used for the *espP* PCR. For primer pair web-7/web-8, the annealing temperature was 60 °C instead of 56 °C.
customized oligonucleotide primers (ARK Scientific). Each base was determined on average three times.

Sequence analysis was conducted with the HUSAR program package (Heidelberg Unix Sequence Analysis Resources, German Cancer Research Centre, Heidelberg, Germany) as well as with the DNASIS program (Hitachi Software).

**Colonel-blot hybridization.** This was performed using a digoxigenin labelling and detection kit (Boehringer Mannheim). Probes were prepared by random labelling of purified espP, katP or EHEC-blyA PCR products, which were obtained using DNA of STEC O157:H7 strain EDL933 as the template.

Bacteria were picked onto Luria agar plates and incubated for 16 h at 37 °C. The plates were cooled at 4 °C for 30 min. Sheets of nylon membrane (Zeta-Probe GT; Bio-Rad) were placed onto the plates for 1 min. Cells were subsequently lysed by alkali as follows. The membrane was laid onto a Whatman 3MM paper soaked with 0.5 M NaOH and 1.5 M NaCl and incubated for 20 min at room temperature. Then the filter was dried on Whatman paper at room temperature for 5 min. The procedure above was repeated twice, then the filter was placed onto a Whatman paper soaked with a solution of 0.5 M NaOH and 0.2% Triton X-100 and incubated for 15 min. After that, the filter was neutralized using 1 M Tris/HCl, pH 7.5, 1 M NaCl. Finally, the membrane was washed for 5 min in 2× SSC (300 mM NaCl, 30 mM sodium citrate), air-dried and baked for 2 h at 80 °C. Hybridization was carried out under stringent conditions according to the manufacturer's instructions. Each strain was tested at least twice.

### RESULTS

**Prevalence of espP in STEC of various serotypes**

To investigate the prevalence of espP sequences in a large number of STEC isolates, we developed a PCR strategy for detection of the espP gene, PCR primers esp-A and esp-B were designed to amplify a 1830 bp fragment of the espP gene (Fig. 1). To ensure a high specificity in the PCR, the primer-binding sites were selected from regions where there was a low level of similarity between the espP gene and the homologous genes espC from EPEC (Stein et al., 1996), espA from Shigella flexneri (Benjelloun-Touimi et al., 1995) and tsh of avian pathogenic *E. coli* (Provenve & Curtiss, 1994).

We first tested the procedure on *E. coli* K-12 laboratory strains harbouring the cloned espP gene (DH5α/pB9 and DH5α/pB9-5; Brudner et al., 1997) and negative control strains DH5α, HB101 and DH5α/pK18 (vector control). The STEC O157:H7 reference strain EDL933 and its plasmid-cured derivative EDL933-cu were also included in this evaluation. As expected, strains DH5α/pB9, DH5α/pB9-5 and EDL933, which harbour the espP gene, gave a positive PCR reaction, whereas the control strains were all negative in the espP PCR.

Next, we examined 156 STEC strains and 45 strains of other diarrhoeagenic *E. coli* (EPEC, EAEC, ETEC and EIEC) using espP PCR (see Table 1 and Methods). Of the 83 O157 strains, 73% possessed the espP gene. It was of interest that when these strains were subdivided according to their capability to ferment sorbitol there was a close correlation with the occurrence of the espP gene: 96% of the classical non-sorbitol-fermenting (NSF) STEC O157:H7/H7 strains possessed the espP gene, whereas the 20 strains of sorbitol-fermenting (SF) STEC O157:H7 tested were all negative in the espP PCR. The occurrence of the espP gene was less common

<table>
<thead>
<tr>
<th>Pathogroup/serotype</th>
<th>Total</th>
<th>espP positive (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>STEC O157:H7/H7</td>
<td>63</td>
<td>61 (96%)*</td>
</tr>
<tr>
<td>NSF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STEC O157:H7</td>
<td>20</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>SF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STEC O26:H11/H7</td>
<td>20</td>
<td>12 (60%)</td>
</tr>
<tr>
<td>SF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STEC O111:H2/H7</td>
<td>8</td>
<td>1 (12%)</td>
</tr>
<tr>
<td>SF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STEC O103:H2</td>
<td>11</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>SF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STEC of other serotypes†</td>
<td>34</td>
<td>15 (44%)</td>
</tr>
<tr>
<td>EPEC</td>
<td>15</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>ETEC</td>
<td>5</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>EIEC</td>
<td>5</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>EAEC</td>
<td>20</td>
<td>0 (0%)</td>
</tr>
</tbody>
</table>

* The two strains negative by espP PCR (3010/96 and 5899/96) hybridized with the espP probe. Strain 5899/96 gave a PCR product of 3-1 kbp instead of the expected product of 1-8 kbp (see text for further comments).


A and esp-B were designed to amplify a 1830 bp fragment of the espP gene (Fig. 1). To ensure a high specificity in the PCR, the primer-binding sites were selected from regions where there was a low level of similarity between the espP gene and the homologous genes espC from EPEC (Stein et al., 1996), espA from Shigella flexneri (Benjelloun-Touimi et al., 1995) and tsh of avian pathogenic *E. coli* (Provenve & Curtiss, 1994).

We first tested the procedure on *E. coli* K-12 laboratory strains harbouring the cloned espP gene (DH5α/pB9 and DH5α/pB9-5; Brudner et al., 1997) and negative control strains DH5α, HB101 and DH5α/pK18 (vector control). The STEC O157:H7 reference strain EDL933 and its plasmid-cured derivative EDL933-cu were also included in this evaluation. As expected, strains DH5α/pB9, DH5α/pB9-5 and EDL933, which harbour the espP gene, gave a positive PCR reaction, whereas the control strains were all negative in the espP PCR.

Next, we examined 156 STEC strains and 45 strains of other diarrhoeagenic *E. coli* (EPEC, EAEC, ETEC and EIEC) using espP PCR (see Table 1 and Methods). Of the 83 O157 strains, 73% possessed the espP gene. It was of interest that when these strains were subdivided according to their capability to ferment sorbitol there was a close correlation with the occurrence of the espP gene: 96% of the classical non-sorbitol-fermenting (NSF) STEC O157:H7/H7 strains possessed the espP gene, whereas the 20 strains of sorbitol-fermenting (SF) STEC O157:H7 tested were all negative in the espP PCR. The occurrence of the espP gene was less common

<table>
<thead>
<tr>
<th>Pathogroup/serotype</th>
<th>Total</th>
<th>espP positive (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>STEC O157:H7/H7</td>
<td>63</td>
<td>61 (96%)*</td>
</tr>
<tr>
<td>NSF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STEC O157:H7</td>
<td>20</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>SF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STEC O26:H11/H7</td>
<td>20</td>
<td>12 (60%)</td>
</tr>
<tr>
<td>SF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STEC O111:H2/H7</td>
<td>8</td>
<td>1 (12%)</td>
</tr>
<tr>
<td>SF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STEC O103:H2</td>
<td>11</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>SF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STEC of other serotypes†</td>
<td>34</td>
<td>15 (44%)</td>
</tr>
<tr>
<td>EPEC</td>
<td>15</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>ETEC</td>
<td>5</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>EIEC</td>
<td>5</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>EAEC</td>
<td>20</td>
<td>0 (0%)</td>
</tr>
</tbody>
</table>

* The two strains negative by espP PCR (3010/96 and 5899/96) hybridized with the espP probe. Strain 5899/96 gave a PCR product of 3-1 kbp instead of the expected product of 1-8 kbp (see text for further comments).


Fig. 1. Variation of the espP gene (white box) in STEC strains of serotype O157:H7. Strain 3010/96 exhibited a partial deletion of the gene. The espP gene of strain 5899/96 is interrupted by insertion of IS1203 (hatched box). Binding sites of the PCR primers esp-A and esp-B and the IS target-site duplication in strain 5899/96 are shown. Sequence positions in base pairs are given with reference to the start codon of espP.
Table 2. Alul/SstI RFLP of the espP PCR product

<table>
<thead>
<tr>
<th>Serotype</th>
<th>No. of strains of restriction type:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
</tr>
<tr>
<td>O157:H7/H+ (NSF)</td>
<td>9</td>
</tr>
<tr>
<td>O26:H11/H-</td>
<td>0</td>
</tr>
<tr>
<td>O145:H+</td>
<td>0</td>
</tr>
<tr>
<td>O111:H+</td>
<td>0</td>
</tr>
<tr>
<td>O22:H8</td>
<td>0</td>
</tr>
<tr>
<td>O69:H-</td>
<td>0</td>
</tr>
<tr>
<td>Orough:H11</td>
<td>0</td>
</tr>
<tr>
<td>O156:H27</td>
<td>0</td>
</tr>
<tr>
<td>O55:H-</td>
<td>0</td>
</tr>
<tr>
<td>O52:H19</td>
<td>0</td>
</tr>
<tr>
<td>O8:H19</td>
<td>0</td>
</tr>
<tr>
<td>ONT:H14/H-</td>
<td>0</td>
</tr>
<tr>
<td>O98:H8</td>
<td>0</td>
</tr>
</tbody>
</table>

To exclude the possibility that the PCR failed to detect the genes due to minor sequence variations in the primer-binding sites, we investigated all strains by colony-blot hybridization using an espP-specific probe. With the exception of STEC O157:H7 strains 3010/96 and 5899/96, which were positive by hybridization but negative in the espP PCR, hybridization of all the strains mentioned above gave the same results as those obtained by PCR.

Deletions and insertions in espP genes of NSF STEC O157:H7

Strains 3010/96 and 5899/96 were examined in detail in order to explain the conflicting results obtained using espP PCR and colony-blot hybridization. Strain 3010/96 was negative in the espP PCR but exhibited hybridization with the espP probe. Strain 5899/96, however, was positive using hybridization, but the product of the espP PCR was approximately 3-1 kbp in length instead of 1-8 kbp. To investigate the nature of the variation in these unusual strains, we hybridized their plasmid DNA with the espP-specific probe and cloned the hybridizing BamHI restriction fragments of 7 kbp (3010/96) and 9-9 kbp (5899/96), into pK18.

Restriction analysis and nucleotide sequencing of the recombinant plasmids revealed that strain 3010/96 had a deletion within the espP gene covering position -138 to 1359 with reference to the ATG start codon of espP. Since the deleted DNA fragment contained the primer-binding site for primer esp-A, this deletion is the cause of the negative PCR reaction (see Fig. 1).

Strain 5899/96, however, was shown to carry a copy of insertion sequence IS1203 (Paton & Paton, 1994), introduced into the espP gene at position 793. This insertion has apparently been caused by a regular transposition event, since we detected direct repeats of the sequence GTTC at the ends of the insertion.
replicating the target-site duplication generated by transposition (Fig. 1).

**RFLP of espP genes**

To investigate the diversity of the espP locus across a broad spectrum of STEC serotypes, we analysed espP PCR products obtained from 36 strains using restriction enzyme digestion with AccI, DraI, EcoRV, AluI and SstI. The subset of strains was representative for all those serotypes in which we had detected the espP gene. Restriction with DraI revealed the same pattern for all strains tested. In contrast, restriction with SstI revealed two different patterns, one of which corresponded to that predicted by the nucleotide sequence of espP derived from strain EDL933 (EMBL accession number X97542; no SstI site within the PCR product), the other to that of the STEC O26 pssA sequence (Djafari et al., 1997; see Fig. 2b). We could identify seven distinct restriction patterns using AluI (restriction types A–G, see Fig. 2). Restriction type E can be discriminated from the similar type A by the presence of one SstI restriction site. All NSF O157:H7/H- isolates showed restriction type A (Table 2). This pattern was not seen in any of the other serotypes analysed. A second restriction pattern, type B, was present in all STEC isolates of serogroup O26 tested and also in two strains of serotype O145:H- and one O111:H- strain. Like restriction type B, types C and E also included strains of different serotypes but no serotype contained more than one restriction type of the espP PCR product (Table 2). Restriction analysis using AccI and EcoRV revealed the same pattern for all strains tested, with the exception of strains of serotype ONT:H14/H- (data not shown).

---

**Table 3. Linkage of the plasmid genes espP (e), katP (k) and EHEC-hlyA (h) in STEC strains of various serotypes**

The number of strains exhibiting the respective marker combination is shown in relation to the serotype.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>e+ k+ h+</td>
<td>63</td>
<td>0</td>
<td>13</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>e+ k+ h-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>e+ k- h+</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>e+ k- h-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>e- k+ h+</td>
<td>0</td>
<td>20</td>
<td>7</td>
<td>5</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>e- k- h+</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>e- k+ h-</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Total</td>
<td>63</td>
<td>20</td>
<td>20</td>
<td>8</td>
<td>11</td>
<td>34</td>
</tr>
</tbody>
</table>

---

**Fig. 3.** Restriction maps of the espP region in pO157 of O157:H7 strain EDL933 (Brunder et al., 1997) and the pssA region in the plasmid of O26:H7 strain 413/89-1 (Djafari et al., 1997). The highly homologous espP and pssA genes, including their promoter and terminator regions, are shown in black. Adjacent sequences with homology to I5s are shown as hatched boxes, together with their inverted terminal repeats (black triangles) and putative transcription direction (small-headed arrows). The black 'lollipops' indicate the hairpin-loop structures of the rho-independent terminators of the espP/pssA genes. The katP gene present in strain EDL933 is depicted as a white box. Large-headed arrows indicate binding sites of the oligonucleotide primers used to investigate the sequences adjacent to espP/pssA. Restriction enzyme sites shown: A, AccI; B, BamHI; E, EcoRI; H, HindIII; K, Kpnl; S, Sall; Sm, Smal; Sh, SphI; St, SstI.
W. BRUNDER and OTHERS

Table 4. Southern hybridization of plasmid DNA with probes specific for espP, katP and EHEC-hlyA

<table>
<thead>
<tr>
<th>Serogroup</th>
<th>Genotype</th>
<th>Strain</th>
<th>espP (kbp)*</th>
<th>katP (kbp)†</th>
<th>hlyA (kbp)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSF O157</td>
<td>e⁺ k⁺ h⁺</td>
<td>3958/95</td>
<td>8.6</td>
<td>9.7</td>
<td>12</td>
</tr>
<tr>
<td>SF O157</td>
<td>e⁺ k⁺ h⁺</td>
<td>3072/96</td>
<td>ND</td>
<td>ND</td>
<td>15</td>
</tr>
<tr>
<td>O26</td>
<td>e⁺ k⁺ h⁺</td>
<td>7662/96</td>
<td>13</td>
<td>4.3</td>
<td>18</td>
</tr>
<tr>
<td>O111</td>
<td>e⁺ k⁺ h⁺</td>
<td>6061/96</td>
<td>ND</td>
<td>ND</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>e⁺ k⁺ h⁺</td>
<td>4993/96</td>
<td>9.8</td>
<td>27</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>e⁺ k⁺ h⁺</td>
<td>94127588</td>
<td>ND</td>
<td>8.2</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>e⁺ k⁺ h⁺</td>
<td>6037/96</td>
<td>ND</td>
<td>ND</td>
<td>13</td>
</tr>
<tr>
<td>O103</td>
<td>e⁺ k⁺ h⁺</td>
<td>5714/96</td>
<td>ND</td>
<td>5.5</td>
<td>12</td>
</tr>
<tr>
<td>O145</td>
<td>e⁺ k⁺ h⁺</td>
<td>1858/96</td>
<td>ND</td>
<td>ND</td>
<td>20</td>
</tr>
<tr>
<td>O103</td>
<td>e⁺ k⁺ h⁺</td>
<td>5157/96I</td>
<td>9</td>
<td>4.4</td>
<td>12</td>
</tr>
<tr>
<td>O145</td>
<td>e⁺ k⁺ h⁺</td>
<td>3985/96</td>
<td>9</td>
<td>ND</td>
<td>20</td>
</tr>
<tr>
<td>O145</td>
<td>e⁺ k⁺ h⁺</td>
<td>0071/96</td>
<td>9.8</td>
<td>ND</td>
<td>28</td>
</tr>
<tr>
<td>O145</td>
<td>e⁺ k⁺ h⁺</td>
<td>1867/96</td>
<td>ND</td>
<td>30</td>
<td>ND</td>
</tr>
<tr>
<td>O145</td>
<td>e⁺ k⁺ h⁺</td>
<td>6481/96</td>
<td>ND</td>
<td>ND</td>
<td>24</td>
</tr>
<tr>
<td>O145</td>
<td>e⁺ k⁺ h⁺</td>
<td>7137/95</td>
<td>Ø</td>
<td>ND</td>
<td>Ø</td>
</tr>
<tr>
<td>O145</td>
<td>e⁺ k⁺ h⁺</td>
<td>6705/95</td>
<td>6.2</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>O55</td>
<td>e⁺ k⁺ h⁺</td>
<td>3379/96</td>
<td>25</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>O25</td>
<td>e⁺ k⁺ h⁺</td>
<td>ED44</td>
<td>ND</td>
<td>ND</td>
<td>9</td>
</tr>
<tr>
<td>O156</td>
<td>e⁺ k⁺ h⁺</td>
<td>4817/96</td>
<td>ND</td>
<td>6</td>
<td>ND</td>
</tr>
<tr>
<td>O156</td>
<td>e⁺ k⁺ h⁺</td>
<td>6561/95</td>
<td>ND</td>
<td>ND</td>
<td>Ø</td>
</tr>
<tr>
<td>O22</td>
<td>e⁺ k⁺ h⁺</td>
<td>6781/95</td>
<td>3.9</td>
<td>ND</td>
<td>20</td>
</tr>
<tr>
<td>O32</td>
<td>e⁺ k⁺ h⁺</td>
<td>4084</td>
<td>25</td>
<td>ND</td>
<td>10</td>
</tr>
<tr>
<td>O52</td>
<td>e⁺ k⁺ h⁺</td>
<td>ED99</td>
<td>9</td>
<td>ND</td>
<td>14</td>
</tr>
<tr>
<td>O69</td>
<td>e⁺ k⁺ h⁺</td>
<td>4014/95</td>
<td>24</td>
<td>ND</td>
<td>9.5</td>
</tr>
<tr>
<td>O8</td>
<td>e⁺ k⁺ h⁺</td>
<td>3651/96</td>
<td>Ø</td>
<td>ND</td>
<td>24</td>
</tr>
<tr>
<td>O98</td>
<td>e⁺ k⁺ h⁺</td>
<td>89/94</td>
<td>30</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>O129</td>
<td>e⁺ k⁺ h⁺</td>
<td>3378/96</td>
<td>ND</td>
<td>6</td>
<td>ND</td>
</tr>
<tr>
<td>O104</td>
<td>e⁺ k⁺ h⁺</td>
<td>4823/96</td>
<td>ND</td>
<td>ND</td>
<td>9</td>
</tr>
<tr>
<td>O113</td>
<td>e⁺ k⁺ h⁺</td>
<td>4395/96</td>
<td>ND</td>
<td>ND</td>
<td>9</td>
</tr>
<tr>
<td>O128</td>
<td>e⁺ k⁺ h⁺</td>
<td>3560/96</td>
<td>ND</td>
<td>ND</td>
<td>23</td>
</tr>
</tbody>
</table>

*DNA restricted with BamHI.
†DNA restricted with SmaI.

Variability at the espP boundaries

In addition to the espP gene itself, we analysed the DNA regions adjacent to the gene. Comparing the nucleotide sequences upstream and downstream of espP and pssA, we found that the region of high homology extended from 138 bp upstream of the ORF to 254 bp downstream of the stop codon. This region included the putative promoter sequences and also the rho-independent terminator described (Brunder et al., 1997; Djafari et al., 1997; see Fig. 3). Beyond this nearly identical DNA section, the two sequences are completely different. Whereas the espP gene is flanked by incomplete insertion sequences (IS8) homologous to IS1203 and iso-IS1 (Brunder et al., 1997), the pssA gene is flanked by two identical elements with inverted orientation and composed of remnants of IS600 and IS911 (Djafari et al., 1997). A complete IS1203-like sequence and a further IS600 remnant occur further downstream on pssA (Fig. 3). We conclude therefore, that in the large plasmid of the STEC O26: H⁻ strain and the STEC O157: H7 strain EDL933 the almost identical genes pssA and espP, together with their regulatory sequences, are inserted at completely different sites.

To investigate the espP boundaries in a greater number of strains, several PCR primers were constructed (see Fig. 3). Using primer pairs web-5/web-6 and web-7/web-8, the expected PCR products were obtained for 10/10 O157: H7/H⁻ strains tested, whereas none of the non-O157 strains gave a PCR product. It is concluded, therefore, that the large plasmids of NSF STEC O157: H7/H⁻ are uniform in the DNA regions adjacent to espP. Another primer pair (web-9/web-6) was constructed based on the sequence upstream of the pssA gene (see Fig. 3). Using this primer pair, PCR products of
the expected length were obtained with 8/11 of the O26 strains tested and also with 2/2 O145:H- and 1/1 O111:H- strains. The other espP-positive serotypes were all negative by PCR using the three primer pairs, indicating that the espP gene of these strains is located within a DNA region different from that described for STEC O157 and STEC O26. Further studies are necessary in order to elucidate the nature of the espP boundaries in these serotypes.

Linkage of espP with other pO157 genes

In a further step we investigated the relationship between the presence of the espP gene in STEC and the presence of katP and the EHEC-haemolysin gene EHEC-blyA, which are known to be also located on plasmid pO157 of STEC O157:H7 strain EDL933 (Schmidt et al., 1995b; Brunder et al., 1996). The genes were also detected by PCR and colony-blot hybridization. Table 3 shows the frequency of all possible combinations of espP (e), katP (k) and EHEC-blyA (h) in STEC strains of specific serotypes. Whereas all NSF STEC O157:H7/H- harboured all three genes, the SF O157:H- strains tested were all negative for espP and katP but positive for EHEC-blyA (by PCR). This is a notable difference between the two groups in addition to the difference in the ability to ferment sorbitol. In STEC strains of serogroup O26, the combinations e-k-h+ and e-k-h occurred, whereas in strains of serotype O103:H2, the predominant combinations were e-k-h+ and e-k-h+ (Table 3). In STEC of other serotypes, the three genes were found in almost all possible combinations.

To prove that all markers are located on plasmids, we performed Southern-blot hybridization of plasmid DNA using probes specific for espP, katP and EHEC-blyA. A representative strain for each marker combination and serogroup was investigated. As shown in Table 4, the genes were detectable on plasmids in all strains with the exception of strain 3651/96 (O8:H19), 7137/95 (Orough:H11) and 6561/95 (O25:H14). In strain 3651/96, EHEC-blyA was detected on a 24 kbp plasmid fragment whereas the espP gene, which is also present, was not detectable on plasmid DNA. It seems likely that in the three cases, the genes (espP and/or EHEC-blyA) are located on the bacterial chromosome. All STEC strains harboured large plasmids, including the nine strains found to be negative for the markers tested. Taken together, the findings clearly demonstrate that STEC large plasmids are a heterogeneous group of genetic elements.

DISCUSSION

Studies on putative virulence factors of STEC encoded by their plasmids have been hitherto confined to a few prototype strains and plasmids. The prototypes have been of serotype O157:H7 because of the prominence of this serotype in outbreaks of STEC infections (Karch et al., 1987, 1998; Tzipori et al., 1987; Toth et al., 1990; Fratamico et al., 1993). In view of the increasing prevalence of non-O157 serotypes in human STEC diseases (Johnson et al., 1996; Caprioli et al., 1997), however, there is now a need for comprehensive data on the virulence properties of these STEC strains. Moreover, virulence genes, including those encoded by plasmids, are used as targets in diagnostic methods and strain subtyping (Nataro & Kaper, 1998). The present study showed that plasmids of STEC, other than the classical serotype O157:H7, possess high variability.

Initially, we analysed the prevalence of espP in a large variety of STEC serotypes and other diarrhoeagenic E. coli and observed a high specificity for STEC among all diarrhoeagenic E. coli pathogens. However, the gene was not detected in about 40% of STEC strains tested. The presence or absence of espP was, in part, correlated with the serotype, but there were also serotypes (e.g. O26:H11/H-) in which only some strains possessed the gene. Some differences in STEC plasmids are to be expected when considering this single marker. Therefore, we included other plasmid markers (katP and EHEC-blyA) and analysed the correlation between the presence of these two additional markers and that of espP. It became apparent that the gene composition is highly variable in STEC plasmids. In contrast to the uniformity present in the group of NSF STEC O157:H7/H-, other serotypes showed nearly all possible combinations of these genes. The fact that plasmids were also detected in the nine STEC strains lacking the three genes indicates that the absence of these genes is not necessarily an indication of a plasmid-less STEC strain.

The SF STEC O157:H- are of particular interest. Strains of this group were isolated in Germany and the Czech Republic from patients suffering from HUS and diarrhoea (Karch et al., 1990; Gunzer et al., 1992; Bielaszewska et al., 1998). In contrast to STEC O157:H7, they ferment sorbitol and exhibit β-glucuronidase activity (Gunzer et al., 1992) and, as shown in the present study, the plasmids of this clone are quite distinct from those of NSF STEC O157:H7/H-. Whereas the haemolysin gene EHEC-blyA is present in both, the plasmids of SF O157:H- strains lack the espP and katP genes. Since NSF E. coli O157:H7 and SF O157:H- are both thought to be derived from a common EPEC-like O55:H7 ancestor (Feng et al., 1998), one can hypothesize that they have acquired different plasmids after their evolutionary split into two lineages. However, it is also possible that one common plasmid, harboured by the ancestor and carrying the EHEC-blyA, espP and katP genes, has lost the espP and katP markers in the SF O157:H- clone. The third possibility is that the ancestral plasmid has acquired the two markers only in the NSF O157:H7 lineage, possibly enhancing the survival and spread of this clone.

One important question is how the acquisition or loss of virulence genes like espP can occur during the evolution of new pathogens. Although there is little known about the source of the espP gene, there are several homologous genes present in E. coli pathogroups and other enteric pathogens, for example the espC gene of EPEC (Stein et al., 1996), the tsh gene of avian pathogenic E. coli
The espP gene itself also shows variability between strains of different serotypes. As we have shown using Aul and SstI restriction of the PCR products, there are at least seven restriction types among espP genes of STEC from the 16 espP-possessing serotypes investigated. These sequence variations, some of which are restricted to one serogroup only, may trace the evolution of the gene within distinct phylogenetic STEC groups. Interestingly, strains of serogroups O26, O111 and O145 are similar, not only in regard to their espP RFLP but also in regard to their espP boundaries. On the other hand, there are other strains within the O26 serogroup showing identical espP-RFLP patterns but different sequences adjacent to the espP gene. However, none of the serotypes tested possessed espP genes of more than one restriction type. Strains with espP restriction type A, B or D also possessed the chromosomal eae gene, whereas those of restriction types C, E, F and G were eae negative (data not shown).

Based on possession of the eae gene and according to the results of DNA hybridization studies, Boerlin et al. (1998) classified STEC plasmids into two categories, type I and type II. They hypothesized that type I plasmids were present in STEC before divergence of the serotypes comprising the eae-positive phylogenetic lineage. Our studies however, show that these type I plasmids are not as uniform as supposed by Boerlin et al., who investigated the EHEC-hlyA gene and the sequences upstream of the EHEC-haemolysin operon. Based on our investigations on espP and katP genes, it is evident that high diversity also exists in the type I plasmids of eae-positive STEC strains. This not only applies to the combinations of the investigated plasmid markers, where marked variations were found in eae-positive strains of serogroups O157, O26 and O103, but also to the DNA regions adjacent to the espP gene.

In summary, our study showed that the large plasmids of STEC are not as uniform as previously supposed. Further efforts are necessary to elucidate the mechanisms of plasmid variability and to determine how variations in the virulence-gene composition influence the pathogenicity of individual STEC strains.

ACKNOWLEDGEMENTS

This work was supported by a grant from the Deutsche Forschungsgemeinschaft (SFB 479). We thank Stefanie Amersbach for excellent technical assistance and Jürgen Scheef for critical reading of the manuscript.

REFERENCES


plasmid of enterohaemorrhagic *Escherichia coli* O157:H7. 
*Microbiology* 142, 907–914.


Received 4 September 1998; revised 10 December 1998; accepted 20 January 1999.