Transfer Region of pO113 from Enterohemorrhagic *Escherichia coli*: Similarity with R64 and Identification of a Novel Plasmid-Encoded Autotransporter, EpeA

Denisse L. Leyton, Joan Sloan, Rebecca E. Hill, Steven Doughty, and Elizabeth L. Hartland*

Bacterial Pathogenesis Research Group, Department of Microbiology, School of Biomedical Sciences,
Monash University, Clayton 3800, Victoria, Australia

Received 3 April 2003/Returned for modification 20 May 2003/Accepted 15 July 2003

Enterohemorrhagic *Escherichia coli* (EHEC) is a prominent, food-borne cause of diarrhea, bloody diarrhea, and the hemolytic uremic syndrome in industrialized countries. Most strains of EHEC carry the locus for enterocyte effacement (LEE) pathogenicity island, but a proportion of isolates from patients with severe disease do not carry LEE and very little is known about virulence factors in these organisms. LEE-negative strains of EHEC typically express Shiga toxin 2 and carry a large plasmid that encodes the production of EHEC hemolysin. In this study, we determined the nucleotide sequence of the transfer region of pO113, the large hemolysin plasmid from LEE-negative EHEC O113:H21 (EH41). This 63.9-kb region showed a high degree of similarity with the transfer region of R64, and pO113 was capable of self-transmission at low frequencies. Unlike R64 and the related *dot/icm* system of *Legionella pneumophila*, however, pO113 was unable to mobilize RSF1010. In addition, the pO113 transfer region encoded a novel high-molecular-weight serine protease autotransporter of *Enterobacteriaceae* (SPATE) protein, termed EpeA. Like other SPATEs, EpeA exhibited protease activity and mucinase activity, but expression was not associated with a cytopathic effect on epithelial cells. Analysis of a second high-molecular-weight secreted protein revealed that pO113 also encodes EspP, a cytopathic SPATE identified previously in EHEC O157:H7. The nucleotide sequences encoding the predicted β -domains of *espP* and *epeA* were identical and also shared significant homology with a third SPATE protein, EspI. Both *espP* and *epeA* were detected in several LEE-negative clinical isolates of EHEC and thus may contribute to the pathogenesis of this subset of EHEC.

Strains of enterohemorrhagic *Escherichia coli* (EHEC) comprise a subset of Shiga toxin (Stx)-producing *E. coli* (STEC) and are a common cause of hemorrhagic colitis (HC) and the hemolytic uremic syndrome (HUS). Although the expression of Stx is a defining feature of virulence, strains of EHEC also harbor a large plasmid encoding EHEC hemolysin (Ehx) and/or a chromosomal pathogenicity island, termed the locus for enterocyte effacement (LEE) (5, 13). Strains of EHEC carrying LEE belong to a group of bacterial pathogens that use a distinctive mechanism of colonization characterized by intimate bacterial adherence and the formation of attaching and effacing (A/E) lesions on the host intestinal mucosa. The most common serotype of EHEC associated with outbreaks and sporadic disease worldwide is O157:H7 (29). EHEC O157:H7 carries LEE and a large hemolysin plasmid, pO157, that is highly conserved among O157:H7 isolates (3, 36). The complete nucleotide sequence of pO157 has revealed an F-like plasmid with 100 predicted open reading frames (ORFs), many encoding putative virulence determinants (5). One of these is a cytopathic, secreted protease, EspP, which is a member of the serine protease autotransporter of *Enterobacteriaceae* (SPATE) family of exported proteins (4). Functional analysis has shown that EspP cleaves pepsin A and human coagulation

factor V and that the protein induces a cytopathic effect on Vero cells.

Although the majority of EHEC strains isolated from patients are LEE-positive A/E pathogens, some serotypes of EHEC do not carry LEE and are not A/E pathogens. These strains have been termed LEE-negative EHEC or STEC and have been regularly associated with sporadic cases and small outbreaks of severe disease, namely HC and HUS (10, 12, 33). Although LEE-negative EHEC serotypes were represented in the original description of EHEC pathogens, the virulence mechanisms of this subset of EHEC have been largely overlooked (11, 23). In the absence of LEE, little is known about the way in which these strains colonize the human intestine, although full adherence undoubtedly requires both plasmid-borne and chromosomal factors (10, 32). Clinical isolates of LEE-negative EHEC typically express Shiga toxin type 2 and also harbor a large plasmid that encodes EHEC hemolysin (12). However, restriction fragment length polymorphism analysis of the *ehxA* gene from LEE-negative EHEC strains has shown that these plasmids comprise an evolutionarily distinct group when compared with the similarly sized plasmids of LEE-positive EHEC (2). Recently, two novel regions of pO113 from LEE-negative EHEC O113:H21 were characterized. One of these regions encoded an autoagglutinating adhesin, Saa, that was associated with the ability of EHEC O113:H21 (98NK2) to adhere to epithelial cells (32). The second region of pO113 comprised a novel type IV pilus gene cluster with predicted amino acid homology to the thin pilus encoded by the self-transmissible IncI plasmid, R64 (25). The type IV pilus

* Corresponding author. Mailing address: Department of Microbiology, Monash University, Clayton 3800, Victoria, Australia. Phone: (61) 3 9905 4323. Fax: (61) 3 9905 4811. E-mail: Liz.Hartland@med.monash.edu.au.

TABLE 1. Bacterial strains and plasmids used in this study

Strain or plasmid	Characteristic(s)	Source or reference
<i>E. coli</i> strains		
EH41	Wild-type EHEC O113:H21 HUS isolate	10
EH41c	pO113-cured strain of EH41	This study
EDL933	Wild-type EHEC O157:H7	34
LT101	Rifampin-resistant derivative of <i>E. coli</i> HB101	31
XL-1 Blue	<i>supE44 hsdR17 recA1 endA1 gyrA46 thi relA1 lac- F' [p roAB+ lacIq lacZ ΔM15 Tn10]</i> (Tet ^r)	Stratagene
Plasmids		
pO113	Large hemolysin plasmid from EHEC O113:H21	This study; 32
R64	Self-transmissible plasmid from <i>Salmonella enterica</i> serovar Typhimurium	26
pMMB67EH	Ampicillin-resistant derivative of RSF1010	15
pCR-Script	High-copy-number cloning vector (Amp ^r)	Stratagene
pKD46	Expression vector for phage λ Red recombinase (γ, β, <i>exo</i>) induced by L-arabinose, thermosensitive Amp ^r	6
p13g	pCR-Script carrying 9.1-kb fragment of pO113 containing <i>epeA</i>	This study
p13g <i>Xho</i>	<i>Xho</i> I fragment deletion derivative of p13g lacking <i>epeA</i>	This study
pCACTUS	Temperature-sensitive suicide plasmid which encodes chloramphenicol resistance and carries <i>sacB</i>	40
pCC1FOS	Copy control cosmid cloning vector	Epicentre

of R64 is required for efficient plasmid transfer in liquid matings, where the pilus promotes donor-recipient contact. Recipient specificity is determined by the minor pilin subunit, PilV, which is under the control of a site-specific shufflon (22). A site-specific recombinase encoded by the shufflon, Rci, rearranges the C-terminal segment of PilV to yield seven alternative adhesins (17). The *pil* genes encoding the thin pilus of R64 are located within a 54-kb *tra* region that encodes one of the most complex DNA transfer systems described to date (26). This region comprises at least 49 genes, 10 of which share similarity with the *dot/icm* secretion system of *Legionella pneumophila* (26).

In this study, we present the complete nucleotide sequence of the transfer region of pO113 and show that components of this region share a high degree of amino acid homology with R64 *tra* gene products. pO113 was self-transmissible at low frequencies when compared with R64 but, unlike R64, pO113 could not mobilize RSF1010 (38). In addition, we found that the pO113 transfer region encoded a novel SPATE protein, EpeA, that exhibited in vitro mucinolytic activity.

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions. The bacterial strains and plasmids used in this study are listed in Table 1. Wild-type strains of pathogenic *E. coli* used to determine the prevalence of *epeA* and *espP* were obtained from the culture collection of Roy Robins-Browne (University of Melbourne) (30). EHEC O113:H21 (EH41) was isolated from a child with HUS in New South Wales, Australia, and expresses Stx2 and EHEC hemolysin (10). R64 was obtained from Derek Pickard (Imperial College London). *E. coli* strains were grown in either Luria-Bertani (LB) broth or serum-free tissue culture medium (Dulbecco's modified Eagle's medium [DMEM]; GIBCO Invitrogen). Broth cultures were incubated aerobically at 37°C to logarithmic phase or overnight. *E. coli* carrying the thermosensitive suicide vector pCACTUS was cultured at 30°C (32, 40). For the induction of λ Red recombinase from plasmid pKD46, 10 mM arabinose was added to LB broth and electrocompetent cells were prepared by growth at 30°C (6). Where necessary, the growth medium was supplemented with 100 μg of ampicillin/ml, 100 μg of kanamycin/ml, 300 μg of rifampin/ml, or 30 μg of chloramphenicol/ml (Sigma Aldrich, St. Louis, Mo.).

DNA and electrophoresis techniques. Standard techniques were employed for all recombinant DNA manipulations, Southern blotting, and electrophoresis procedures, including sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (35). DNA-modifying enzymes (Promega, Madison, Wis.) were used according to the manufacturer's instructions. Chromosomal DNA was iso-

lated using the DNeasy tissue kit (Qiagen, Hilden, Germany), and plasmid DNA was isolated using the Qiaprep Spin miniprep kit (Qiagen).

To prepare a pO113 plasmid library, plasmid DNA from EHEC strain O113:H21 (EH41) was extracted and digested with *Bam*HI. The resulting fragments were cloned randomly into the *Bam*HI site of pCR-Script, and the nucleotide sequences of both strands were determined using M13-based sequencing primers or custom-made oligonucleotides. The plasmid clone, p13g, derived from this library carries a 9.1-kb fragment of pO113 that includes the entire *epeA* gene. p13g*Xho* is a derivative of p13g where a *Xho*I fragment containing *epeA* has been deleted by restriction enzyme digestion and religation.

To obtain larger fragments of pO113 for sequence analysis, a random genomic library was created in the Copy Control pCC1FOS cosmid vector according to the manufacturer's instructions (Epicentre, Madison, Wis.). pO113 *tra* clones were then selected by dot blot hybridization with digoxigenin (DIG)-labeled probes generated by PCR (Roche Molecular Biochemicals, Basel, Switzerland), and nucleotide sequencing was performed directly on purified pCC1FOS DNA.

Nucleotide sequencing and analysis. Plasmid or cosmid DNA was prepared for sequence analysis using a PRISM Ready Reaction DyeDeoxy Terminator cycle sequencing kit (Applied Biosystems, Foster City, Calif.). The nucleotide sequence of samples was determined by automated DNA sequencing using an Applied Biosystems model 373A DNA sequencing system. DNA sequences were assembled using Sequencer 3.1.1 (Gene Codes Corp., Ann Arbor, Mich.). GeneMark.hmm for prokaryotes (<http://opal.biology.gatech.edu/GeneMark>) was used to identify putative genes, and BLAST programs (<http://www.ncbi.nlm.nih.gov:80/BLAST>) were used to determine nucleotide and amino acid homologies with sequences in GenBank. The cellular localizations of putative proteins were predicted using the PSORT programs (<http://psort.ims.u-tokyo.ac.jp>).

Transfer of pO113 and RSF1010. Liquid and surface bacterial conjugations were performed as described previously (24). For pO113 transfer experiments, EH41 carrying pO113*ehx:km* was used as the donor strain. For mobilization experiments, the donor strains EHEC O113:H21 (EH41) and *E. coli* carrying R64 were transformed with an ampicillin-resistant derivative of RSF1010, pMMB67EH (15). EH41c and *E. coli* LT101 (31) acted as rifampin-resistant recipient strains. Log-phase cultures of donor cells were mixed with overnight cultures of recipient cells, and the mixtures were incubated at 37°C for 90 min before plating onto LB agar containing kanamycin and rifampin for selection of pO113*ehx:km* transfer or ampicillin and rifampin for selection of pMMB67EH mobilization. Transfer frequencies were expressed as the ratio of transconjugants to donor cells.

Construction of a pO113-cured derivative and *epeA* and *ehx* mutants of EHEC O113:H21 (EH41). Curing of pO113 was achieved by recombination of the thermosensitive vector pCACTUS into pO113 (26). Briefly, a 1.3-kb segment of the *ehx* operon was cloned into pCACTUS, which was then electroporated into EHEC O113:H21 (EH41). Recombination of pCACTUS:*ehx* with pO113 was selected by growth at 37°C on Luria agar supplemented with chloramphenicol. The resulting colonies were EHEC hemolysin negative on EHEC hemolysin agar plates, confirming the interruption of *ehx*. Curing of the recombined pO113:CACTUS plasmid was achieved by overnight growth of this strain at 43°C in LB

TABLE 2. Sequence of PCR primers used in this study

Gene	Primer sequences	Annealing temp (°C)	Reference
<i>epeA</i>	EpeAF: 5'-CACCTGTAGAAATCTTA-3' EpeAR: 5'-CTGAATAAATCCAGCCC-3'	46	This study
<i>espP</i>	EspPF: 5'-AAACAGCAGGCACTTGAACG-3' EspPR: 5'-GGAGTCGTCAGTCAGTAGAT-3'	56	4
<i>beta</i>	BetaF: 5'-GGCTCTGCCAGTGGTGG-3' BetaR: 5'-GTTGTACTTACCAAAGG-3'	42	This study
<i>ehx</i>	EhxF: 5'-GGAACCGCTGAAAATGTAGG-3' EhxR: 5'-ACTGGTCGTCTCCCTGTCC-3'	50	This study
<i>pilS</i>	PilSF: 5'-TCCAGTGACAACCAGCGC-3' PilSR: 5'-TTAGCTGTTGGTTTCCAG-3'	52	This study
<i>tra</i>	TraF: 5'-AAACAGCAGGCACTTGAACG-3' TraR: 5'-AAACAGCAGGCACTTGAACG-3'	50	This study
<i>O113lps</i>	O113F: 5'-AGCGTTTCTGACATATGGAGTG-3' O113R: 5'-GTGTTAGTATCAAAAGAGGCTCC-3'	50	33

broth supplemented with 6% sucrose. Bacteria were then diluted and spread onto EHEC hemolysin agar. The vast majority of colonies were EHEC hemolysin positive, indicating the excision of pCACTUS:*ehx*. However, 12 EHEC hemolysin-negative colonies were identified and tested for the presence of pO113 by PCR using primers designed for the pO113-encoded genes *ehx*, *pilS*, *traB*, and *epeA*. The resulting pO113-cured strain was designated EH41c and confirmed to be a derivative of EHEC O113:H21 (EH41) by PCR using O113 lipopolysaccharide-specific primers.

To construct an *epeA* mutant of EHEC O113:H21 (EH41), we used the λ Red recombinase system to introduce a kanamycin resistance gene into *epeA* located on pO113. Briefly, the temperature-sensitive replicon pKD46 was introduced into EH41, and electrocompetent cells were prepared from bacteria grown at 30°C with 10 mM arabinose to induce Red recombinase expression as described previously (6). A kanamycin cassette was cloned into the *Pst*I site located 1.3 kb downstream of the putative start codon of *epeA*, and the interrupted *epeA* gene fragment was amplified by PCR and introduced into EH41 containing pKD46 by electroporation. Successful interruption of *epeA* was confirmed by PCR analysis and was evident by an increase in molecular weight of 1,185 bp corresponding to the insertion of the kanamycin cassette (data not shown) and loss of EpeA from the supernatant of EH41 cultures. Using a similar approach, we ligated the kanamycin cassette into a *Pst*I site 310 bp from the start codon of *ehxA* and introduced the mutation onto pO113 using Red recombinase. Successful integration of the kanamycin resistance marker was confirmed by PCR and the absence of hemolysis on EHEC hemolysin agar.

PCR procedures and generation of DNA probes. The PCR primer pairs used in this study are listed in Table 2. PCR amplification was performed on 500 ng of template DNA with approximately 1 μ g of each primer PCR per 100- μ l reaction mixture. Amplification of *epeA*, the β -domain of *epeA/espP* (*beta*), *ehx*, *traB*, and *pilS*, was performed under the following conditions: 2 min at 94°C and 30 cycles of 1 min at 94°C, 50 s at appropriate annealing temperature (Table 2), and 1 min at 72°C, followed by 5 min at 72°C. PCR amplification of *espP* was performed under the following conditions: 5 min at 94°C and 30 cycles of 30 s at 94°C, 1 min at 56°C, and 2.5 min at 72°C, followed by 5 min at 72°C as described previously (4), using the primers EspPF and EspPR. PCR-generated DNA products were examined by agarose gel electrophoresis, and where necessary DIG-dUTP was incorporated into the PCR according to the manufacturer's instructions (Roche) to generate DIG-labeled DNA probes.

Preparation of culture supernatant proteins. For SDS-PAGE, secreted proteins were precipitated from bacterial culture supernatants. Briefly, overnight bacterial broth cultures were diluted 1:20 in 10 ml of prewarmed DMEM and grown to mid-log phase. The bacteria were pelleted by centrifugation (10,000 $\times$ g, 4°C, 10 min), and the supernatants were filtered through 0.22- μ m-pore-size filters (Millipore, Bedford, Mass.). A 10% (wt/vol) final concentration of trichloroacetic acid (TCA; BDH Laboratory Supplies, Poole, Dorset, England) was used to precipitate the proteins, which were then pelleted by centrifugation (10,000 $\times$ g, 4°C, 45 min). The protein pellets were washed twice with 10 ml of 100%

(vol/vol) methanol (BDH Laboratory Supplies), pelleted as before, dried for 1 h at 37°C, and finally resuspended in 2 $\times$ SDS sample buffer.

For functional assays, proteins were concentrated from culture supernatants. Overnight broth cultures were diluted 1:20 into 800 ml of prewarmed LB and grown to mid-log phase. Bacteria were then pelleted by centrifugation (10,000 $\times$ g, 4°C, 10 min), and supernatants were filtered through 0.22- μ m-pore-size filters (Millipore). The clarified supernatants were then concentrated 400-fold by ultrafiltration through PM10, PM30, or XM50A membranes (Amicon Corp., Lexington, Mass.). All manipulations were performed at 4°C.

N-terminal amino acid sequencing. Automated Edman degradation was used to determine the N-terminal amino acid sequences of secreted proteins. The proteins were separated by SDS-10% PAGE and blotted onto a polyvinylidene difluoride sequencing-grade membrane (Perkin-Elmer Inc., Wellesley, Mass.). After blotting, proteins were visualized by staining with Coomassie brilliant blue (Bio-Rad Laboratories), and the appropriate bands were excised and sequenced using an Applied Biosystems 476A amino acid sequencer.

Assays for protease and mucinase activity. Assays for protease activity with swine pepsin A (Roche) were performed as described previously with a minor variation (37). Briefly, 5 μ g of concentrated supernatant was mixed with 3 μ g of pepsin A, resuspended in phosphate-buffered saline (PBS) to a final volume of 20 μ l, and incubated overnight at 37°C. Protein mixtures were then separated by SDS-PAGE, transferred to nitrocellulose, and detected with rabbit antipepsin antibodies (provided by I. van Driel). Protease inhibition assays were performed in the presence of 10 μ g of phenylmethylsulfonyl fluoride (PMSF)/ml. Zymogram analysis of gelatin proteolysis was adapted from a previous method (41). In this assay, 8 μ g of the preparation of secreted proteins was electrophoresed into a 10% nondenaturing PAGE gel containing 2 μ g of gelatin (BDH Biochemical)/ml. After electrophoresis, the gel was incubated for 24 h in 40 mM Tris, 2 mM CaCl₂ (pH 8.0) at 37°C, and the gel was then stained with 0.125% (wt/vol) Coomassie brilliant blue and destained in 40% (vol/vol) methanol-10% (vol/vol) acetic acid. Proteases from *Dichelobacter nodosus* were used as a positive control (28).

The lysis of bovine submaxillary mucin was assessed as described previously (18). Briefly, equivalent preparations of concentrated supernatant proteins were added to wells bored into agar containing 1.5% (wt/vol) agarose, 1.0% (wt/vol) glucose, and 0.5% (wt/vol) mucin in Luria broth. Following incubation at 37°C for 24 h, the agar was stained with 0.1% (wt/vol) amido black in 3.5 M acetic acid. Lysis of mucin was observed as a halo of clearing around the inoculation point.

Assays for cytotoxicity. Assays for cytotoxicity were carried out as described by Guyer et al. (16) using cultured HeLa cells and CHO-K1 cells. Cells were stained with Giemsa (Sigma) and observed for morphological changes by bright-field microscopy using a Leica DM-RB HC fluorescence phase-contrast microscope.

Immunodetection of EpeA. To obtain antibodies specific for EpeA, a region encoding the C-terminal 300 amino acids of the passenger domain of EpeA was amplified by PCR and cloned into pET28a to generate a His₆-tagged fusion, His-EpasC. Rabbits were immunized four times with 100 μ g of purified recom-

FIG. 1. Schematic representation of the 63.9-kb transfer region of pO113 and comparison with R64. Grey arrows represent ORFs shared by pO113 and R64. White arrows represent ORFs unique to pO113, and black arrows indicate ORFs present in R64 and not pO113. Arrows indicate the direction of gene transcription. The scale shown is for pO113, and the R64 gene names are provided below the corresponding ORFs. The dotted line represents the absence of an equivalent R64 DNA sequence.

TABLE 3. Properties of putative pO113 *tra* gene products

Gene	No. of amino acids	Mass (kDa)	pI	Predicted location ^a	Homologue	Property	Origin	% Identity/similarity
<i>traB</i>	213	24.6	9.5	C	TraB	Regulator of transfer	R64	65/79
<i>traC</i>	221	25.0	7.8	C	TraC	Regulator of transfer	R64	36/51
<i>ygiJ</i>	206	22.4	5.3	IM	YgiJ	Oxidoreductase	<i>E. coli</i> K-12	57/72
<i>B3051</i>	564	62.2	5.0	OM	B3051	Membrane protein	<i>E. coli</i> K-12	81/92
<i>pilI</i>	80	9.1	9.2	C	PilI		R64	32/51
<i>pilL</i>	356	38.8	10.5	IM/OM	PilL	Lipoprotein	R64	42/55
<i>pilM</i>	145	16.3	7.0	IM	PilM	Pilus biogenesis	R64	29/53
<i>pilN</i>	539	57.5	7.7	OM	PilN	Secretin	R64	39/60
<i>pilO</i>	431	47.7	6.0	IM	PilO	Pilus biogenesis	R64	22/38
<i>pilP</i>	150	16.3	6.3	P/OM	PilP	Pilus biogenesis	R64	27/42
<i>pilQ</i>	502	56.0	6.4	C	PilQ	Nucleotide-binding protein	R64	43/58
<i>pilR</i>	366	41.0	9.4	IM	PilR	Pilus biogenesis	R64	28/49
<i>pilS</i>	147	15.9	9.1	OM	PilS	Type IV prepilin	R64	25/45
<i>pilT</i>	157	17.8	9.9	OM	PilT	Transglycosylase	R64	52/66
<i>pilU</i>	211	23.1	9.6	IM	PilU	Prepilin peptidase	R64	25/42
<i>pilV</i>	454	49.0	8.4	OM	PilV	Type IV prepilin	R64	41/56
<i>orf1</i>	38	4.2	12.3	OM	None			
<i>orf2</i>	55	6.2	5.7	C	None			
<i>artA</i>	106	12.1	9.3	IM	ArtA		F-plasmid	34/60
<i>traE</i>	273	31.2	8.9	C	TraE		R64	56/72
<i>traF</i>	184	19.6	5.5	P/OM	TraF (first 165 aa) ^b		R64	71/81
<i>traF'</i>	221	24.1	5.1	C	TraF (last 220 aa)		R64	78/91
<i>traH</i>	150	16.5	8.9	IM/OM	TraH		R64	64/77
<i>traI</i>	278	31.7	10.1	IM/OM	TraI	Lipoprotein	R64	78/86
<i>traJ</i>	382	42.5	6.4	IM	TraJ	Nucleotide-binding protein	R64	78/89
<i>traK</i>	96	11.5	9.8	IM	TraK		R64	71/86
<i>sogL</i>	1,338	146.1	4.8	C	SogL	DNA primase	R64	52/65
<i>traL</i>	116	12.6	6.6	P/OM	TraL		R64	73/84
<i>traM</i>	231	25.5	9.0	IM	TraM		R64	70/83
<i>traN</i>	324	34.2	5.6	P/OM	TraN		R64	76/85
<i>traO</i>	443	47.1	7.2	IM	TraO		R64	58/71
<i>traP</i>	237	26.2	7.8	IM	TraP		R64	47/61
<i>traQ</i>	175	18.1	8.0	IM	TraQ		R64	76/86
<i>traR</i>	132	7.1	9.6	IM/OM	TraR		R64	54/74
<i>traS</i>	83	9.1	6.3	IM	TraS		R64	67/79
<i>traT</i>	272	31.1	9.6	C	TraT		R64	60/78
<i>orf3</i>	49	5.5	4.7	C	None			
<i>orf4</i>	127	14.8	10.5	P/OM	Psy0721		<i>P. aeruginosa</i>	56/76
<i>traU</i>	1,014	114.0	6.5	C	TraU	Nucleotide-binding protein	R64	77/89
<i>traV</i>	205	23.4	9.5	C	TraV		R64	66/81
<i>traW</i>	401	42.9	5.9	IM/OM	TraW	Lipoprotein	R64	67/78
<i>traX</i>	189	22.0	10.5	IM	TraX		R64	40/58
<i>traY</i>	721	77.3	6.7	IM	TraY	Integral membrane protein	R64	65/78
<i>excA</i>	216	25.1	9.1	IM	ExcA	Surface exclusion	R64	48/66
<i>orf5</i>	320	36.3	9.9	C	RpoS		<i>R. solanacearum</i>	31/57
<i>orf6</i>	224	26.2	8.9	C	Gmet1519		<i>G. metallireducens</i>	28/45
<i>pndC</i>	94	10.5	9.0	C	PndC		R64	82/86
<i>orf7</i>	120	13.4	7.1	C	C3606		UPEC	68/78
<i>orf8</i>	411	46.6	6.0	P/OM	YigX		UPEC	52/67
<i>orf9</i>	207	23.7	9.6	C	Z1195		EHEC O157:H7	64/74
<i>yigB</i>	103	11.6	9.8	P/OM	YigB		R100	56/68
<i>epeA</i>	1,359	147.3	6.4	OM	EspI		STEC	58/71
<i>L7021</i>	36	3.9	5.3	C	L7021		pO157	96/100
<i>orf10</i>	63	7.3	9.9	C	C3605		UPEC	71/82
<i>orf11</i>	119	13.3	6.6	C	C3604		UPEC	40/59
<i>orf12</i>	121	14.1	6.5	C	C3603		UPEC	68/75
<i>orf13</i>	63	7.0	7.8	IM	C3602		UPEC	85/91
<i>orf14</i>	103	12.0	9.9	C	R8	Neurotensin receptor	UPEC	57/65
<i>trbA</i>	423	48.6	9.2	IM	TrbA		R64	67/78
<i>trbB</i>	372	40.3	5.8	IM	TrbB		R64	50/63
<i>trbC</i>	767	87.1	7.6	IM	TrbC		R64	75/84
<i>orf15</i>	100	11.5	7.8	C	C5174	Iron-regulated protein	UPEC	66/81
<i>L7092</i>	149	16.7	9.0	IM/OM	L7092		pO157	98/100
<i>nikB</i>	903	104.5	7.2	C	NikB	Relaxosome complex	pO157	94/94
<i>nikA</i>	139	16.1	9.4	C	NikA	Relaxosome complex	R64	60/75

^a C, cytoplasm; IM, innermembrane; OM, outer membrane; P, periplasm.^b aa, amino acids.

FIG. 2. Schematic representation of a 9.1-kb *Bam*HI fragment of pO113 containing *epeA*. The white arrow represents *epeA* (4,080 bp), and the dark arrows represent other putative ORFs encoded by this fragment. Cleavage positions for *Kpn*I, *Sac*I, *Bam*HI, and *Xho*I are indicated, and the direction of *lacZ* transcription from pCR-Script is shown. Structure of EpeA is indicated below the 9.1-kb fragment and shows the N- and C-terminal cleavage sites, serine protease, and P-loop motifs and their corresponding amino acid sequences and positions. The predicted molecular mass of each domain is indicated.

binant His-EpasC at 3-week intervals. The resulting antiserum was then absorbed three times with the *epeA* mutant of EH41. To determine the effect of temperature and pH on EpeA secretion, strain EH41 was grown in 25 ml of LB broth to an optical density at 600 nm (OD_{600}) of 1.0 at differing pH and temperatures, and supernatant proteins were precipitated with 10% (wt/vol) TCA. Supernatant proteins were then separated by SDS-PAGE, and EpeA was detected by immunoblotting with antibodies raised against His-EpasC.

Nucleotide sequence accession number. The nucleotide sequence presented in this study has been assigned GenBank accession no. AY258503.

RESULTS

Genetic organization of the pO113 transfer region. The nucleotide sequence of the 63.9-kb pO113 transfer region revealed 65 putative ORFs, including 41 homologues of R64 transfer genes (Table 3). The overall genetic organization of the pO113 transfer region and comparison with R64 are shown in Fig. 1. Overall, the G+C content of the pO113 transfer region was 51%, which is similar to that of the *E. coli* K-12 genome (50.8%). Amino acid identity between pO113 products and their R64 homologues varied from 25 to 82% (Table 3). Major differences between the R64 and pO113 transfer regions were the absence of *traA*, *traD*, and *traG* in pO113 and a frameshift in *traF*. However, these *tra* genes are not required

for R64 transfer (26). In addition, pO113 lacked *nuc*, *pilJ*, and *pilK* and the *pilV* shuffle. Although *pilV* was present, there were no inverted repeat regions characteristic of the R64 shuffle and, importantly, the site-specific recombinase *rci*, which is responsible for *pilV* rearrangements in R64, was missing. These sequence data agreed with a previous *pil*_{O113} sequence determined from a different strain of EHEC O113:H21, which has been shown to code for type IV pili (38). Instead of the R64 *nuc* gene, pO113 carried *yigB*, a gene encoding a 15.6-kDa putative nuclease similar to YigB from R100 (8). The pO113 transfer region also encoded several ORFs not present in R64. Two of these, *ygiJ* and *B3051*, were located within the *pil* region and encoded a putative oxidoreductase and outer membrane protein, respectively. *ygiJ* and *B3051* are present in EHEC O157:H7 and *E. coli* K-12 but are located on the chromosome in these organisms. The additional 10.7-kb region located between *pndC* and *trbA* that was not present in R64 comprised 11 predicted ORFs, including a gene encoding a putative SPATE protein that we termed EpeA for EHEC plasmid-encoded autotransporter (Fig. 2; Table 3). Analysis of the region immediately upstream of the 4.1-kb *epeA* gene revealed a homologue of Z1195 (Orf9), a hypothetical protein with unknown

TABLE 4. Conjugal transfer of pO113*ehx:km* and mobilization of RSF1010(pMMB67EH) by pO113 and R64

Plasmid transferred	Transfer or mobilization frequency (%) ^a			
	Filter matings		Liquid matings	
	LT101 ^b	EH41c	LT101	EH41c
pO113 <i>ehx:km</i>	0.00006	0.00009	0.00004	0.0001
pMMB67EH				
by R64 ^c	14.4	0.066	3.4	0.0088
by pO113	<0.00001	<0.00001	<0.00001	<0.00001

^a Transfer or mobilization frequency of pO113*ehx:km* and pMMB67EH from donor cells. Results are expressed as the percentage of transconjugants relative to donor cells.

^b Recipient strains *E. coli* LT101 or pO113-cured EHEC O113:H21 (EH41c).

^c Donor plasmids R64 or pO113 for mobilization of pMMB67EH.

function in EHEC O157:H7 EDL933, and investigation of the region immediately downstream from *epeA* revealed a small ORF identical to *L7021* encoded by pO157 (5). The 35-amino-acid predicted product of *L7021* has no known homologues but is located adjacent to *espP* on pO157.

pO113 transfer and mobilization of RSF1010 by pO113 and R64. Previous work has suggested that pO113 is self-transmissible (38). In this study we confirmed that pO113 is capable of conjugal transfer but at a much lower frequency (~0.00005%) than that reported for R64 (~4.6%) (Table 4) (26). In addition, there was no significant difference in the transfer frequency of pO113 using *E. coli* LT101 or EH41c as recipients or between liquid or filter conjugations. Due to the high degree of similarity between pO113 and R64, we also tested the ability of pO113 to mobilize RSF1010; however, we were unable to detect mobilization of an RSF1010 derivative, pMMB67EH, by pO113 in this study (Table 4).

Identification of *epeA* and homology with other SPATE proteins. The *epeA* gene located in the pO113 transfer region was predicted to encode a 1,359-amino-acid precursor protein with a molecular mass of 147 kDa. EpeA was comparable in length and exhibited a significant degree of amino acid similarity to several SPATE proteins, including Pic, Tsh, SepA, and EspI (Table 5) (1, 18, 37, 39). Sequence analysis showed that like other SPATE proteins, EpeA comprised an extended N-terminal leader peptide signal sequence, a passenger or functional α -domain, and a C-terminal β -domain for which the cleavage site was predicted to lie between residues N₁₀₈₂ and N₁₀₈₃ (Fig. 2). The C-terminal β -domain also exhibited a conserved 3-amino-acid motif, YSF, essential for outer membrane localization of Hap, a SPATE protein from *Haemophilus influenzae* (20, 21). In addition, the β -domain contained an ATP/GTP-binding site motif (P-loop), A₁₁₅₂DVFSGKT, which is

carried by many nucleotide-binding proteins (19, 20). Importantly, the putative passenger domain of EpeA possessed a potential serine protease motif, G₂₅₄DSGSP, which was conserved in sequence and location in comparison to other SPATE proteins (Fig. 2) (19). Interestingly, while the nucleotide sequence of the passenger domain was unique and showed no homology to other nucleotide sequences present in GenBank, the β -domain exhibited 99% nucleotide homology to the β -domain of *espP* (GenBank accession no. AF074613) and 91% nucleotide homology to the β -domain of *espI* (accession no. AJ278144). Other genetic homologues that share this β -domain include *aaaA* from prophage P-EibA (87% homology; accession no. AF151091) and also *pet* from enteroaggregative *E. coli* and *sat* from uropathogenic *E. coli* (83% homology; accession numbers AF056581 and AF289092, respectively). Overall, the G+C content of the *epeA* gene was 41.9%, markedly lower than that of the entire pO113 transfer region and the *E. coli* K-12 chromosome (50.8%).

Expression of EpeA by EHEC O113:H21. To determine if *epeA* was functional, we examined the secreted protein profile of EHEC O113:H21 (EH41) for the presence of a high-molecular-weight secreted protein. Analysis of the secreted protein fraction revealed two dominant high-molecular-weight proteins of 104 and 112 kDa which appeared to be common to several strains of EHEC O113:H21 and other serotypes of LEE-negative EHEC (Fig. 3). N-terminal amino acid sequencing was performed on both proteins. The amino acid sequence, (Y/S/H/K)TVDAEIP, obtained for the 112-kDa protein corresponded to the first 8 amino acids predicted for the secreted form of EpeA and was located 52 amino acids from the predicted start of the precursor protein. The second sequence obtained for the 104-kDa protein, (S/K)QMDISNF, exhibited 100% identity with the first 8 amino acids of the mature, secreted form of EspP, a product of the large hemolysin plasmid, pO157, carried by EHEC O157:H7.

Functional analysis of EpeA. Based on the presence of a conserved serine protease motif and homology to a number of SPATE proteins, all of which exhibit protease activity, we postulated that EpeA also possessed proteolytic activity. To test the protease activity of EpeA, we concentrated supernatant proteins from two *E. coli* strains expressing EpeA, namely wild-type EHEC O113:H21 (EH41) and *E. coli* XL1-Blue harboring the plasmid p13g, which carries a 9.1-kb fragment of pO113 including a complete copy of *epeA* (Fig. 2 and 4). As a negative control, we constructed an *epeA* mutant of EHEC O113:H21 (EH41) by the insertion of a kanamycin resistance marker into a unique *Pst*I site within the coding region of *epeA*. This mutant was designated EH41*epeA*, and analysis of the secreted protein profile of the mutant showed that interruption

TABLE 5. Amino acid sequence identity and similarity of EpeA with its closest homologues

Protein	Accession no.	Major reported function	No. of amino acids	% Identity (% similarity) with:			
				Entire protein	Leader sequence	Passenger domain	β -Domain
EspI	CAC39286	Protease	1,363	58 (71)	91 (96)	47 (63)	96 (97)
SepA	CAC05786	Tissue invasion	1,364	50 (65)	58 (75)	43 (60)	79 (91)
Pic	AAK00464	Mucinase	1,373	49 (64)	64 (84)	41 (57)	80 (90)
Tsh	I54632	Hemagglutinin	1,377	42 (58)	52 (75)	37 (54)	62 (78)

FIG. 3. SDS-PAGE gel showing the dominant high-molecular-mass proteins, EpeA and EspP, in secreted protein fractions from LEE-negative strains of EHEC. Lane 1, EHEC O157:H7 (EDL933) (LEE positive); lane 2, EHEC O113:H21 (EH41); lane 3, EHEC O113:H21 (EH53); lane 4, EHEC O113:H21 (EH71); lane 5, EHEC O116:H21 (EH42); lane 6, EHEC O130:H11 (EH43); lane 7, EHEC O1:H7 (EH69). Bacteria were grown to mid-log phase, and supernatant proteins were precipitated with TCA and resolved by an SDS-12.5% PAGE gel. Arrows indicate protein bands corresponding to EpeA and EspP. Arrowheads indicate LEE-encoded proteins.

of *epeA* resulted in loss of the 112-kDa protein from the culture supernatant (Fig. 4). In addition, we constructed a control for p13g by the deletion of a *XhoI* fragment containing *epeA*, and this plasmid was termed p13g*Xho*. *E. coli* XL1-Blue (p13g*Xho*) also failed to secrete a 112-kDa protein into the culture supernatant (Fig. 4).

Concentrated supernatant proteins prepared from all four strains were used to assay protease activity in a pepsin A-based assay as described previously (37). Complete degradation of pepsin A was observed when concentrated supernatant proteins derived from EH41 and *E. coli* XL1-Blue(p13g) were incubated with the substrate (Fig. 5A, lanes 2 and 4). Conversely, EH41*epeA* and *E. coli* XL1-Blue(p13g*Xho*) supernatant proteins did not cleave pepsin A (Fig. 5A, lanes 3 and 5). In addition, preincubation for 1 h at 37°C with 10 µg of PMSF/ml abolished the proteolytic activity of the concentrated supernatants from EH41 and *E. coli* XL1-blue(p13g) cells (data not shown), indicating that protease activity was responsible for the cleavage of pepsin A in this assay. In addition to cleavage of pepsin A, we tested the ability of EpeA to digest gelatine in a zymogram assay. In this system, concentrated supernatant proteins from EH41 and *E. coli* XL1-Blue(p13g) produced extensive zones of clearance indicating the proteolysis of gelatine (Fig. 5B, lanes 2 and 4), whereas protein samples derived from EH41*epeA* and *E. coli* XL1-Blue(p13g*Xho*) showed no protease activity (Fig. 5B, lanes 3 and 5).

Mucinas activity has been reported for the *Shigella flexneri* SPATE protein, Pic, which is closely related to EpeA by amino acid similarity (18). Here we also tested the ability of EpeA to

FIG. 4. Secreted proteins isolated from EHEC O113:H21 (EH41) (lane 1), EH41*epeA* (lane 2), *E. coli* XL1-Blue(p13g) (lane 3), and *E. coli* XL1-Blue(p13g*Xho*) (lane 4). Bacterial strains were grown to mid-log phase, and supernatant proteins were precipitated with TCA and resolved on an SDS-12.5% PAGE gel. The open arrowhead indicates EpeA, and the closed arrowhead indicates EspP. The asterisk indicates a breakdown product of EpeA.

FIG. 5. (A) Immunoblot demonstrating cleavage of swine pepsin A by concentrated supernatant proteins derived from EH41 (lane 2), EH41*epeA* (lane 3), *E. coli* XL1-Blue(p13g) (lane 4), and *E. coli* XL1-Blue(p13g*Xho*) (lane 5). The arrow indicates uncleaved pepsin, which is also shown for reference in lane 1. (B) Gelatinase zymogram analysis of concentrated supernatant proteins derived from EH41 (lane 2), EH41*epeA* (lane 3), *E. coli* XL1-Blue(p13g) (lane 4), and *E. coli* XL1-Blue(p13g*Xho*) (lane 5). Lane 1 shows activity of secreted proteases from *D. nodosus* as a positive control. (C) Mucinolytic activity of EpeA. Agar containing 0.5% bovine submaxillary mucin was inoculated with 5 μ g of supernatant proteins derived from *E. coli* XL1-Blue(p13g) (i) and *E. coli* XL1-Blue(p13g*Xho*) (ii) per ml. Agar was stained with 0.1% amido black.

cleave bovine submaxillary mucin. Concentrated supernatant proteins derived from *E. coli* XL1-Blue(p13g) and *E. coli* XL1-Blue(p13g*Xho*) were added to agar containing 0.5% (wt/vol) bovine submaxillary mucin. Incubation with supernatant pro-

teins from *E. coli* XL1-Blue(p13g) containing EpeA produced a clear zone of lysis around the inoculation site (Fig. 5C, well i), whereas incubation with supernatant proteins from *E. coli* XL1-Blue(p13g*Xho*) failed to produce a zone of clearing in the

FIG. 6. Effect of temperature and pH on EpeA secretion. Western blot analysis of EpeA in whole-cell protein extracts (a lanes) and precipitated supernatant proteins (b lanes) from EH41 cultures grown to an OD_{600} of 1.0 at 24, 30, 37, and 39°C (A) and pH 4.0, 6.0, and 8.0 (B).

agar containing mucin (Fig. 5C, well ii). These results indicated that, similar to Pic, EpeA can lyse bovine submaxillary mucin (18).

Although EpeA is less related by homology to the cytopathic group of SPATE proteins, we nevertheless tested the cytotoxic activity of EpeA against HeLa cells. We were unable, however, to attribute a cytopathic effect to EpeA at a concentration of 25 μ g/ml, which was shown previously to be sufficient for induction of cytotoxicity by other SPATE proteins such as EspP, Sat, and Pet (4, 14, 16).

Effect of temperature and pH on EpeA secretion. To ascertain if temperature and/or pH regulates EpeA secretion as it does for Pic, we monitored the amount of EpeA released into the culture supernatant from EH41 cells grown under different conditions to the same optical density ($OD_{600} = 1.0$). EpeA was detected by immunoblotting using anti-EpeA antibodies raised in rabbits. The results showed that EpeA was expressed maximally at 37 to 39°C and at alkaline pH ($pH \geq 8$) as described for Pic (Fig. 6) (18). Interestingly, EpeA was not detected in whole-cell protein extracts, suggesting that the majority of the protein is released into the culture supernatant (Fig. 6).

***epeA* and *espP* share a common β -domain.** A BLASTN comparison of *espP* and *epeA* suggested that these two genes share a near-identical β -domain. As this level of nucleotide homology has not been reported among the β -domains of SPATE genes previously, we performed a Southern blot analysis on EH41 and the pO113-cured derivative of EH41 (EH41c) to confirm that two copies of the β -domain were present on pO113. Firstly, specific PCR analysis of whole-cell DNA from EH41c showed that *epeA* and *espP* were missing from the cured strain, thus confirming that both genes are located on pO113 (data not shown). For Southern analysis, genomic DNA from EH41 and EH41c was digested with *EcoRI*, which does not cut within the β -domain, and hybridized with a DIG-labeled probe, *beta*, that was specific for the β -domain of *epeA*. The *beta* probe was generated by PCR using p13g as a template and the primers BetaF and BetaR. Unexpectedly, the Southern blotting showed that EH41 carried three copies of the β -domain and that EH41c carried one copy (Fig. 7). This result confirmed that two β -domains were present on pO113, and we

hypothesized that the third copy of the β -domain carried on the chromosome of EH41 was part of a large genomic island described recently in a serogroup O91:H- LEE-negative strain of STEC which encodes EspI (37). Further PCR analysis confirmed that EH41 carried this genomic island inserted at the *selC* locus and suggested that our *beta* probe had hybridized to the β -domain region of an *espI* homologue (data not shown).

Distribution of *epeA* and *espP* among LEE-negative and LEE-positive strains of EHEC. N-terminal amino acid sequencing of two high-molecular-weight secreted proteins iden-

FIG. 7. Southern blot analysis of *EcoRI*-digested genomic DNA derived from EH41c (lane 1) and EH41 (lane 2). The membrane was hybridized with DIG-labeled *beta* probe specific for the β -domain of *epeA*.

TABLE 6. Prevalence of *epeA* and *espP* among different serotypes of EHEC

Serotype	<i>eae</i> ^a	No. of positive strains/no. tested		
		<i>espP</i>	<i>epeA</i>	<i>beta</i>
O157:H7	+	2/2	0/2	2/2
O111:H-	+	1/4	0/4	3/4
O15:H-	+	1/1	0/1	1/1
O26:H11	+	2/2	0/2	2/2
O113:H21	—	11/11	9/11	11/11
O116:H21	—	1/1	1/1	1/1
O130:H11	—	1/1	1/1	1/1
O5:H-	—	0/2	1/2	2/2
O1:H7	—	1/1	1/1	1/1
NT:H7	—	1/1	1/1	1/1
O48:H21	—	1/1	1/1	1/1
O76:H7	—	0/1	0/1	1/1
O128:H2	—	0/3	0/3	3/3
O91:H-	—	2/2	0/2	2/2

^a Marker for LEE pathogenicity island.

tified in this study revealed that EH41 secreted a homologue of EspP. As the presence of EspP in LEE-negative strains of EHEC has not been described previously, we examined the distribution of *espP* and the novel gene *epeA* among strains of EHEC and EPEC using PCR primers specific for their respective passenger domains (Table 6). The bacterial strains collected for testing have been described previously in other studies (10, 12, 30, 32). The majority of LEE-negative EHEC strains (15 of 24 strains tested) possessed the *epeA* gene, while none of the LEE-positive EHEC strains tested carried *epeA* (Table 6). As *epeA* is a novel gene, we also tested its prevalence among isolates of EPEC and other *E. coli* pathogens, including nine strains of EPEC and five isolates each of enteroaggregative *E. coli*, enteroinvasive *E. coli*, enterotoxigenic *E. coli*, and *S. flexneri*; however, all were negative (data not shown). Thus, *epeA* appeared to be specifically associated with LEE-negative isolates of EHEC. Interestingly, although *epeA* was not found in every LEE-negative EHEC serotype tested, it was found in 85% of strains associated with severe disease, i.e., isolates derived from patients with HUS and thrombotic thrombocytopenic purpura and/or HC (12). In contrast to *epeA*, *espP* was found in both LEE-positive strains of EHEC, as reported previously (3), and in the majority of LEE-negative EHEC strains (18 of 24 strains tested) (Table 6). Interestingly, 97% of the EHEC strains tested (32 of 33) were positive for the β -domain (*beta*) of *espP* and *epeA*, while only 1 of 15 EPEC isolates tested (7%) was positive. This serogroup O142:H6 EPEC strain was also positive for *espP*. In addition, *beta* was not detected by PCR in the other *E. coli* pathogens described above, suggesting that a *beta*-specific PCR may be a useful molecular diagnostic test for the detection of EHEC in combination with detection methods for *stx* and/or *ehx* and *eae*.

DISCUSSION

LEE-negative strains of EHEC are a significant cause of HUS and HC in many parts of the world (33). In Australia, EHEC O113:H21 is the second most common serotype isolated from patients with HUS after O111:H- (12). To date, the virulence factors possessed by LEE-negative strains of EHEC

remain poorly characterized, and in particular the determinants that mediate adherence to the gastrointestinal tract in the absence of LEE are unknown. Recently, we described a novel chromosomal fimbrial gene cluster in EHEC O113:H21 that shares predicted similarity to long polar fimbrial components (10). This locus was required for full adherence of EHEC O113:H21 to epithelial cells and thus may play a role in colonization of the host intestine.

Like EHEC O157:H7, LEE-negative strains of EHEC carry a large plasmid that encodes EHEC hemolysin and EspP, although apart from these two factors, the LEE-negative EHEC plasmids appear to be unrelated to pO157. In this study we determined the nucleotide sequence of the transfer region of pO113, the large hemolysin plasmid from EHEC O113:H21 (32). This region shared a high degree of similarity with components of the *tra* region from R64, a self-transmissible plasmid identified originally in *Salmonella enterica* serovar Typhimurium (26). Although in this study we confirmed that pO113 was self-transmissible, the transfer frequency of pO113 was around 10⁵-fold lower than that reported for R64 (26). In addition, unlike R64 and the related *dot/icm* system of *L. pneumophila*, pO113 was unable to mobilize the IncQ plasmid RSF1010. This suggests that the two systems are functionally different and may indicate that some components of the pO113 transfer region are not expressed or functional.

Although gene orientation and order among homologues of the R64 and pO113 transfer regions were largely conserved, we observed several differences, including the presence of an additional 10.7-kb region in pO113 located between *pndC* and *trbA*. Several putative ORFs were present in this region that shared similarity with hypothetical proteins from uropathogenic *E. coli* and EHEC O157:H7. The function and origin of these hypothetical genes in either pathogen is unclear, but together they represent a combination of plasmid-borne and chromosomal factors, suggesting that, like pO157, pO113 has a mosaic structure and has evolved through the acquisition of determinants from a number of sources. In addition, we identified a novel SPATE protein in the additional 10.7-kb region, termed EpeA. EpeA was one of two dominant high-molecular-weight proteins secreted by LEE-negative strains of EHEC that was also identified by N-terminal amino acid sequencing. By amino acid similarity, EpeA was most closely related to EspI, a SPATE protein described recently in a serogroup O91:H- strain of LEE-negative STEC. However, EpeA was also highly related to SPATEs from other pathogens, namely SepA, Pic, and Tsh (1, 18, 37, 39). EpeA possessed protease and mucinase activity but was not able to induce a cytopathic effect on HeLa cells. Like its close relative Pic, EpeA was maximally expressed at body temperature under alkaline conditions, suggesting that the protein may be produced as the pathogen enters the small intestine (18). Although the contribution of EpeA to pathogenesis is unclear, the protein may aid colonization and adherence to the host intestine through mucinolytic activity.

The second high-molecular-weight secreted protein identified in EHEC O113:H21 (EH41) was EspP. EspP is a cytopathic SPATE protein which is encoded by pO157 in EHEC O157:H7 (4, 9). In EHEC O113:H21 (EH41), both *epeA* and *espP* were encoded by pO113, although *espP* was not located within the 63.9-kb transfer region. We identified *epeA* and *espP*

in the majority of LEE-negative EHEC isolates tested, and *espP* was also prevalent in LEE-positive strains of EHEC as reported previously (4). Importantly we have shown that, together with *ehx* and *stx*, *espP* is common to both LEE-negative and LEE-positive clinical isolates of EHEC and thus may represent a good molecular marker for the detection of EHEC in general. In addition, a PCR specific for the β -domain nucleotide sequence common to both *epeA* and *espP* was able to detect 97% of EHEC strains tested in this study (32 of 33), including 7 strains that were PCR negative for the *espP* and *epeA* passenger domains. This may indicate that these isolates carry additional autotransporters, such as EspI, that utilize the β -domain of *epeA* and *espP* but have a unique passenger domain. Indeed, autotransporters are believed to have evolved by domain shuffling where a new passenger domain is linked with a generic β -domain (7, 27).

The nucleotide sequence of pO157 contains remnants of a transfer system that indicate this plasmid was once conjugative. Through evolution, however, the F-plasmid-like transfer region of pO157 has been interrupted and replaced by insertion sequence elements and putative virulence genes (5). Although pO113 has retained much of its transfer region, several additional genes are present within this region that are not found in R64 or other closely related transfer systems. The acquisition of these additional genes and the loss of other components such as the shufflon may reflect the changing role of pO113 from a conjugative plasmid to a stable virulence-associated plasmid like pO157. Overall, the results of this study suggest that the megaplasmids of LEE-negative EHEC have a different evolutionary origin than the plasmids of LEE-positive EHEC strains but nevertheless share at least some virulence determinants, including EHEC hemolysin and EspP. The complete nucleotide sequence of pO113 will help to resolve the genetic relationship between pO113 and pO157 and will aid our understanding of the role of these plasmids in the pathogenesis of EHEC infections.

ACKNOWLEDGMENTS

We gratefully acknowledge the Australian Paediatric Surveillance Unit and Roy Robins-Browne for the provision of EHEC and other *E. coli* strains. R64 was kindly provided by Derek Pickard. We are also indebted to Ian van Driel for the provision of pepsin antibodies used in this study and to Dane Parker for providing *D. nodosus* proteases.

This work was supported in part by funding from the Australian National Health and Medical Research Council, ANZ Charitable Trusts, Ramaciotti Foundations for Biomedical Research, and Monash University. D.L.L. is the recipient of an Australian Postgraduate Award.

REFERENCES

- Benjelloun-Touimi, Z., P. J. Sansonetti, and C. Parsot. 1995. SepA, the major extracellular protein of *Shigella flexneri*: autonomous secretion and involvement in tissue invasion. *Mol. Microbiol.* **17**:123–135.
- Boerlin, P., S. Chen, J. K. Colbourne, R. Johnson, S. De Grandis, and C. Gyles. 1998. Evolution of enterohemorrhagic *Escherichia coli* hemolysin plasmids and the locus for enterocyte effacement in Shiga toxin-producing *E. coli*. *Infect. Immun.* **66**:2553–2561.
- Brunder, W., H. Schmidt, M. Frosch, and H. Karch. 1999. The large plasmids of Shiga-toxin-producing *Escherichia coli* (STEC) are highly variable genetic elements. *Microbiology* **145**:1005–1014.
- Brunder, W., H. Schmidt, and H. Karch. 1997. EspP, a novel extracellular serine protease of enterohaemorrhagic *Escherichia coli* O157:H7 cleaves human coagulation factor V. *Mol. Microbiol.* **24**:767–778.
- Burland, V., Y. Shao, N. T. Perna, G. Plunkett, H. J. Sofia, and F. R. Blattner. 1998. The complete DNA sequence and analysis of the large virulence plasmid of *Escherichia coli* O157:H7. *Nucleic Acids Res.* **26**:4196–4204.
- Datsenko, K. A., and B. L. Wanner. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. USA* **97**:6640–6645.
- Davis, J., A. L. Smith, W. R. Hughes, and M. Golomb. 2001. Evolution of an autotransporter: domain shuffling and lateral transfer from pathogenic *Haemophilus* to *Neisseria*. *J. Bacteriol.* **183**:4626–4635.
- Dempsey, W. B. 1994. Regulation of R100 conjugation requires *traM* in *cis* to *traJ*. *Mol. Microbiol.* **13**:987–1000.
- Djafari, S., F. Ebel, C. Deibel, S. Kramer, M. Hudel, and T. Chakraborty. 1997. Characterization of an exported protease from Shiga toxin-producing *Escherichia coli*. *Mol. Microbiol.* **25**:771–784.
- Doughty, S., J. Sloan, V. Bennett-Wood, M. Robertson, R. M. Robins-Browne, and E. L. Hartland. 2002. Identification of a novel fimbrial gene cluster related to long polar fimbriae in locus of enterocyte effacement-negative strains of enterohemorrhagic *Escherichia coli*. *Infect. Immun.* **70**:6761–6769.
- Dytoc, M. T., A. Ismaili, D. J. Philpott, R. Soni, J. L. Brunton, and P. M. Sherman. 1994. Distinct binding properties of *eaeA*-negative verocytotoxin-producing *Escherichia coli* of serotype O113:H21. *Infect. Immun.* **62**:3494–3505.
- Elliott, E. J., R. M. Robins-Browne, E. V. O'Loughlin, P. Henning, G. G. Hogg, J. Knight, H. Powell, D. Redmond, and V. Bennett-Wood. 2001. Nationwide surveillance of haemolytic uraemic syndrome in Australia: a comparison of endemic cases with those in a single epidemic. *Arch. Dis. Child.* **85**:1–4.
- Elliott, S. J., L. A. Wainwright, T. K. McDaniel, K. G. Jarvis, Y. K. Deng, L. C. Lai, B. P. McNamara, M. S. Donnenberg, and J. B. Kaper. 1998. The complete sequence of the locus of enterocyte effacement (LEE) from enteropathogenic *Escherichia coli* E2348/69. *Mol. Microbiol.* **28**:1–4.
- Eslava, C., F. Navarro-Garcia, J. R. Czeizulin, I. R. Henderson, A. Cravioto, and J. P. Nataro. 1998. Pet, an autotransporter enterotoxin from enteroaggregative *Escherichia coli*. *Infect. Immun.* **66**:3155–3163.
- Furste, J. P., W. Pansegrau, R. Frank, H. Blocker, P. Scholz, M. Bagdasarian, and E. Lanka. 1986. Molecular cloning of the plasmid RP4 primase region in a multi-host-range *tacP* expression vector. *Gene* **48**:119–131.
- Guyer, D. M., I. R. Henderson, J. P. Nataro, and H. L. Mobley. 2000. Identification of Sat, an autotransporter toxin produced by uropathogenic *Escherichia coli*. *Mol. Microbiol.* **38**:53–66.
- Gyohda, A., N. Furuya, N. Kogure, and T. Komano. 2002. Sequence-specific and non-specific binding of the Rci protein to the asymmetric recombination sites of the R64 shufflon. *J. Mol. Biol.* **318**:975–983.
- Henderson, I. R., J. Czeizulin, C. Eslava, F. Noriega, and J. P. Nataro. 1999. Characterization of Pic, a secreted protease of *Shigella flexneri* and enteroaggregative *Escherichia coli*. *Infect. Immun.* **67**:5587–5596.
- Henderson, I. R., F. Navarro-Garcia, and J. P. Nataro. 1998. The great escape: structure and function of the autotransporter proteins. *Trends Microbiol.* **6**:370–378.
- Hendrixson, D. R., M. L. de la Morena, C. Stathopoulos, and J. W. St. Geme III. 1997. Structural determinants of processing and secretion of the *Haemophilus influenzae* Hap protein. *Mol. Microbiol.* **26**:505–518.
- Hendrixson, D. R., and J. W. St. Geme III. 1998. The *Haemophilus influenzae* Hap serine protease promotes adherence and microcolony formation, potentiated by a soluble host protein. *Mol. Cell* **2**:841–850.
- Ishiwa, A., and T. Komano. 2000. The lipopolysaccharide of recipient cells is a specific receptor for PilV proteins, selected by shufflon DNA rearrangement, in liquid matings with donors bearing the R64 plasmid. *Mol. Gen. Genet.* **263**:159–164.
- Karmali, M. A., M. Petric, C. Lim, P. C. Fleming, G. S. Arbus, and H. Lior. 1985. The association between idiopathic hemolytic uremic syndrome and infection by verotoxin-producing *Escherichia coli*. *J. Infect. Dis.* **151**:775–782.
- Kim, S. R., N. Funayama, and T. Komano. 1993. Nucleotide sequence and characterization of the *traABCD* region of Inc11 plasmid R64. *J. Bacteriol.* **175**:5035–5042.
- Kim, S. R., and T. Komano. 1997. The plasmid R64 thin pilus identified as a type IV pilus. *J. Bacteriol.* **179**:3594–3603.
- Komano, T., T. Yoshida, K. Narahara, and N. Furuya. 2000. The transfer region of Inc11 plasmid R64: similarities between R64 *tra* and *Legionella icm/dot* genes. *Mol. Microbiol.* **35**:1348–1359.
- Loveless, B. J., and M. H. Saier, Jr. 1997. A novel family of channel-forming, autotransporting, bacterial virulence factors. *Mol. Membr. Biol.* **14**:113–123.
- Moses, E. K., J. I. Rood, W. K. Yong, and G. G. Riffkin. 1989. Molecular analysis of one of multiple protease-encoding genes from the prototype virulent strain of *Bacteroides nodosus*. *Gene* **77**:219–228.
- Nataro, J. P., and J. B. Kaper. 1998. Diarrheagenic *Escherichia coli*. *Clin. Microbiol. Rev.* **11**:142–201.
- Nicholls, L., T. H. Grant, and R. M. Robins-Browne. 2000. Identification of a novel genetic locus that is required for in vitro adhesion of a clinical isolate of enterohaemorrhagic *Escherichia coli* to epithelial cells. *Mol. Microbiol.* **35**:275–288.
- Palombo, E. A., K. Yusoff, V. A. Stanisich, V. Krishnapillai, and N. S.

- Willettts. 1989. Cloning and genetic analysis of *tra* cistrons of the Tra 2/Tra 3 region of plasmid RP1. *Plasmid* **22**:59–69.
32. Paton, A. W., P. Srimanote, M. C. Woodrow, and J. C. Paton. 2001. Characterization of Saa, a novel autoagglutinating adhesin produced by locus of enterocyte effacement-negative Shiga-toxigenic *Escherichia coli* strains that are virulent for humans. *Infect. Immun.* **69**:6999–7009.
 33. Paton, A. W., M. C. Woodrow, R. M. Doyle, J. A. Lanser, and J. C. Paton. 1999. Molecular characterization of a Shiga toxigenic *Escherichia coli* O113: H21 strain lacking *eae* responsible for a cluster of cases of hemolytic-uremic syndrome. *J. Clin. Microbiol.* **37**:3357–3361.
 34. Perna, N. T., G. Plunkett III, V. Burland, B. Mau, J. D. Glasner, D. J. Rose, G. F. Mayhew, P. S. Evans, J. Gregor, H. A. Kirkpatrick, G. Posfai, J. Hackett, S. Klink, A. Boutin, Y. Shao, L. Miller, E. J. Grotbeck, N. W. Davis, A. Lim, E. T. Dimalanta, K. D. Potamou, J. Apodaca, T. S. Anantharaman, J. Lin, G. Yen, D. C. Schwartz, R. A. Welch, and F. R. Blattner. 2001. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* **409**:529–533.
 35. Sambrook, J., E. F. Fritsch, T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Press, Cold Spring Harbor, N.Y.
 36. Schmidt, H., and H. Karch. 1996. Enterohemolytic phenotypes and genotypes of Shiga toxin-producing *Escherichia coli* O111 strains from patients with diarrhea and hemolytic-uremic syndrome. *J. Clin. Microbiol.* **34**:2364–2367.
 37. Schmidt, H., W. L. Zhang, U. Hemmrich, S. Jelacic, W. Brunder, P. I. Tarr, U. Dobrindt, J. Hacker, and H. Karch. 2001. Identification and characterization of a novel genomic island integrated at *selC* in locus of enterocyte effacement-negative, Shiga toxin-producing *Escherichia coli*. *Infect. Immun.* **69**:6863–6873.
 38. Srimanote, P., A. W. Paton, and J. C. Paton. 2002. Characterization of a novel type IV pilus locus encoded on the large plasmid of locus of enterocyte effacement-negative Shiga-toxigenic *Escherichia coli* strains that are virulent for humans. *Infect. Immun.* **70**:3094–3100.
 39. Stathopoulos, C., D. L. Provenge, and R. Curtiss III. 1999. Characterization of the avian pathogenic *Escherichia coli* hemagglutinin Tsh, a member of the immunoglobulin A protease-type family of autotransporters. *Infect. Immun.* **67**:772–781.
 40. Van den Bosch, L., P. A. Manning, and R. Morona. 1997. Regulation of O-antigen chain length is required for *Shigella flexneri* virulence. *Mol. Microbiol.* **23**:765–775.
 41. Young, A. R., N. Mancuso, E. N. Meeusen, and V. M. Bowles. 2000. Characterization of proteases involved in egg hatching of the sheep blowfly, *Lucilia cuprina*. *Int. J. Parasitol.* **30**:925–932.

Editor: A. D. O'Brien