JOURNAL OF CLINICAL MICROBIOLOGY, June 2003, p. 2341–2347 Vol. 41, No. 6
0095-1137/03/$08.00+0 DOI: 10.1128/JCM.41.6.2341–2347.2003

Distribution of the Secondary Type III Secretion System Locus Found in Enterohemorrhagic *Escherichia coli* O157:H7 Isolates among Shiga Toxin-Producing *E. coli* Strains

Sou-Ichi Makino,[1,2]* Toru Tobe,[3] Hiroshi Asakura,[1] Masahisa Watarai,[1] Testuya Ikeda,[4] Koichi Takeshi,[4] and Chihiro Sasakawa[5]

Department of Applied Veterinary Science, Research Center for Animal Hygiene and Food Safety,[1] Obihiro University of Agriculture and Veterinary Medicine,[2] and Hokkaido Institute of Public Health,[4] Hokkaido, Division of Applied Bacteriology, Graduate School of Medicine, Osaka University, Osaka,[3] and Department of Microbiology and Immunology, Institute of Medical Science, University of Tokyo, Tokyo,[5] Japan

Received 18 October 2002/Returned for modification 20 January 2003/Accepted 18 February 2003

The ability of the complete genome sequence of enterohemorrhagic *Escherichia coli* O157 led to the identification of a 17-kb chromosomal region which contained a type III secretion system gene cluster at min 64.5. This locus contains open reading frames whose amino acid sequences show high degrees of similarity with those of proteins that make up the type III secretion apparatus, which is encoded by the *inv-spa-prg* locus on a *Salmonella* SPI-1 pathogenicity island. This locus was designated ETT2 (*E. coli* type III secretion 2) and consisted of the *epr*, *epa*, and *eiv* genes. ETT2 was found in enteropathogenic *E. coli* strains and also in some non-O157 Shiga toxin-producing *E. coli* (STEC) strains, but most of them contained a truncated portion of ETT2. Most O157 isolates had a complete collection of toxin-encoding genes *eae* and *hlyA* and the ETT2 locus, while most O26 strains had toxin-encoding genes *eae* and *hlyA* genes but an incomplete ETT2 locus. Thus, an intact copy of ETT2 might mark a pathogenic distinction for particular STEC strains. Therefore, the presence of the ETT2 locus can be used for identification of truly pathogenic STEC strains and for molecular fingerprinting of the epidemic strains in humans and animals.

Enterohemorrhagic *Escherichia coli* (EHEC) and Shiga toxin-producing *E. coli* (STEC) are the most common causes of hemorrhagic colitis, bloody diarrhea, and hemolytic-uremic syndrome and also cause food-borne epidemics in humans (19). STEC is distributed widely in the environment (1, 2). It produces cytotoxin as the main virulence factor but can have a histopathologic effect on intestinal epithelial cells, causing attaching-and-effacing (A/E) lesions. Enteropathogenic *E. coli* (EPEC), which is a major cause of infant diarrhea in developing countries, has the same ability to cause A/E lesions. These lesions are characterized by localized destruction of brush border microvilli, intimate attachment of the organism to the host cell membrane, and formation of an actin-rich underlying structure in the host cells (5). All genes necessary for the formation of A/E lesions are located in a pathogenicity island termed the locus for enterocyte effacement (LEE) (18). This locus contains genes for structural components of a secretion apparatus that belongs to the type III secretion system, that is, genes for the adhesin intimin and its translocated receptor, Tir (7).

The type III secretion system, which is found in many gram-negative pathogens, is responsible for secretion and injection of virulence-associated factors into the cytosol of host cells. The type III apparatus comprises approximately 20 proteins, with most of them located in the inner membrane (6, 24). Most inner membrane proteins are homologous to components of the flagellar biosynthesis apparatuses of both gram-negative and gram-positive bacteria (10). In *Salmonella enterica* serovar Typhimurium, two independent type III secretion systems encoded by separate chromosomal loci, SPI-1 and SPI-2, have been found (8, 22). SPI-1 is required for the invasion of mammalian cells, and SPI-2 is required for replication in macrophages and for survival in mice.

The release of the complete genome sequence of EHEC O157:H7 (9) provided us with an opportunity to identify sequences specific for STEC strains. In this study, we report on a type III secretion locus in the STEC chromosome similar to the SPI-1 system. To estimate the roles of this locus in virulence expression in STEC, we show the distribution of this novel locus among various STEC strains, and we also discuss the transmission of STEC from animals to humans.

MATERIALS AND METHODS

Bacterial strains and media. Table 1 lists all STEC strains used in the study. EHEC O157:H7 strain RIMD 0509952 (referred to here as EHEC O157 Sakai) was originally isolated from a patient in an outbreak in Sakai, Japan, in 1996 (14). Four EPEC isolates from humans were used in this study: B171-8 (O111:NM, where NM indicates nonmotile) (20), E2348/69 (O127:H7), 4394-57 (O114:NM), and 1929-55(O126:NM). Some STEC strains were isolated from deer (1), sheep (2), seagulls (2), houseflies (11), and salmon roe (15). Other STEC and EPEC strains are the stock isolates in our laboratory. Bacteria were cultured in Trypticase soy broth (BBL/Becton Dickinson, Cockeysville, Md.) or on Trypticase soy agar (BBL/Becton Dickinson).

Southern hybridization. Chromosomal DNA was extracted from EHEC, EPEC, and *E. coli* K-12 strains as described previously (17) and digested with *Eco*RI. The DNA fragments were separated by electrophoresis in 1% agarose gels and were transferred to Hybond-N+ positively charged nylon membranes (Amersham, Little Chalfont, England). DNA fragments corresponding to three separate regions in ETT2 of EHEC O157:H7 (region A, *eprH-eprK*; region B,

* Corresponding author. Mailing address: Department of Applied Veterinary Science, Obihiro University of Agriculture and Veterinary Medicine, Inada-cho, Obihiro, Hokkaido 080-8555, Japan. Phone and fax: 81-155-49-5386. E-mail: smakino@obihiro.ac.jp.

TABLE 1. STEC strains used in this study

Strain	Origin	Serotype	Toxin	PCR result (band size)					Western blot results for *eaeA*
				L (6.2 kb)	D (2.6 kb)	C (5.5 kb)	*eae*	*hlyA*	
S-1	Deer	O111:H45	2	− (1.9 kb)	−	−	−	+	−
S-2	Deer	O111:H45	2	− (1.9 kb)	−	−	−	+	−
S-3	Deer	O111:H45	2	− (1.9 kb)	−	−	−	+	−
S-4	Deer	Out:H45	2	−	−	−	−	+	−
S-5	Deer	O93:H−	2	−	−	−	C	+	−
S-6	Deer	O96:H−	2	− (1.9 kb)	−	−	C	+	−
S-7	Deer	O96:H−	2	− (1.9 kb)	−	−	C	+	−
S-8	Deer	Out:H25	2	+	+	+	−	−	−
S-9	Deer	O8:H7	2	− (1.9 kb)	−	−	−	−	−
S-10	Deer	O96:H−	2	− (1.9 kb)	−	−	−	−	−
S-11	Deer	O96:H−	2	− (1.9 kb)	−	−	−	−	−
HI-1	Sheep	O153:H25	1	− (<1.0 kb)	−	−	−	−	−
HI-2	Sheep	O165:H−	1, 2	− (2.5 kb)	−	−	C	+	−
HI-5	Sheep	Out:H16	1	−	−	−	−	−	−
HI-7	Sheep	Out:H−	1	+	4 kb	+	−	+	−
HI-11	Sheep	O2:Hut	2	+	3 kb	+	−	+	−
HI-N	Sheep	Out:H−	1,2	+	−	−	−	+	−
K-7	Seagull	O136:H16	2	+	+	− (0.6 kb)	−	+	−
K-10	Seagull	O153:H−	1	+	−	−	−	−	−
fly-11	Fly	ONT:H−	1, 2	+	−	−	−	−	−
F-16	Fly	O157:H7	1, 2	+	+	+	γ	+	+
F-156	Fly	O157:H7	1, 2	+	+	+	γ	+	+
F-158	Fly	O157:H7	1, 2	+	+	+	γ	+	+
F2-1	Fly	O157:H7	−	+	+	+	γ	+	+
Obi-1*	Human	O157:H7	2	+	+	+	γ	+	+
Sakai	Human	O157:H7	1, 2	+	+	+	γ	+	+
T-1*	Human	O157:H−	1	+	+	+	γ	+	+
T-46*	Human	O157:H7	1, 2	+	+	+	γ	+	+
K-159*	Human	O157:H7	2	+	+	+	γ	+	+
K-262*	Human	O157:H7	1, 2	+	+	+	γ	+	+
EC097028*	Human	O157:H7	−	+	+	+	γ	+	+
SR	Salmon roe	O157:H7	1, 2	+	+	+	γ	+	+
AS16	Cattle	O157:H7	2	+	+	−	γ	+	+
AS17	Cattle	O157:H7	2	+	+	+	γ	+	+
AS18	Cattle	O157:H7	2	+	+	−	γ	+	+
AS19	Cattle	O157:H7	2	+	+	−	γ	+	+
AS20	Cattle	O157:H7	1, 2	+	+	−	γ	+	+
KT1	Cattle	O157:H7	2	+	+	−	γ	+	+
HK1	Cattle	O157:H7	2	+	+	+	γ	+	+
MB1	Cattle	O157:H7	1	+	+	+	γ	−	+
TS−	Cattle	O157:H7	2	+	+	+	γ	+	+
IK3	Cattle	O157:H7	2	+	+	+	γ	+	+
KR1	Cattle	O119	1, 2	+	−	−	β	+	+
IK4	Cattle	O119	1	−	−	−	β	+	+
AS1*	Human	O26:H11	1, 2	+	−	−	β	+	+
AS15	Human	O26:H11	1, 2	+	−	−	β	+	+
H4-52	Cattle	O26:H−	1	−	−	−	β	+	+
H4-54	Cattle	O26:H11	1	−	−	−	β	−	+
H5-7	Cattle	O26:H−	1	−	−	−	β	−	+
H5-16*	Cattle	O26:H−	1	−	−	−	β	−	+
H5-26	Cattle	O26:H11	1	+	−	−	β	−	+
H5-160	Cattle	O26:H11	1	−	−	−	β	−	+
H7-10	Cattle	O26:H11	1	−	−	−	β	+	+
H8-55	Cattle	O26:H11	1	+	−	−	β	+	+
H8-87	Cattle	O26:H11	1	+	−	−	β	−	+
IK2	Cattle	O26	1	+	−	−	β	+	+

[a] Symbols and abbreviations: *, seven STEC strains used for Southern hybridization; −, negative; +, positive; C, *eae* common primer positive but nontypeable; γ, *eae*; γ-specific primer positive; β, *eae* β-specific primer positive.

epaO-epaS; and region D, *eivF-eivA*) were amplified by PCR with primer sets specific for each region (Table 2) and were labeled with biotin by using the BrightStar Psoralen-Biotin nonisotopic labeling kit (Ambion, Austin, Tex.). Signals were detected by using the BrightStar BioDetect nonisotopic detection kit (Ambion).

PCR. Table 2 and Table 3 show the oligonucleotide primers and conditions used for PCR. All primers (20 pmol each) were mixed with TaKaRa Ex *Taq* (Takara Bio Inc., Shiga, Japan), and the PCR was performed in a volume of 25 μl in a model 9600 thermal cycler (Applied Biosystems Japan Ltd., Tokyo, Japan).

Western blotting. Western blotting was performed as described previously (16) with rabbit polyclonal antibody raised against intimin-γ.

Nucleotide sequence accession number. The DNA sequence of EPEC strain B171-8 (O111:NM) has been submitted to GenBank and can be found under accession no. AB052736.

TABLE 2. Primers used to detect genes in five separate regions in the ETT2 locus[a]

Region	Genes in ETT2	Primer sequence (5′ to 3′)
A	*eprH-eprK*	TAGAATGGCCCTATACTTCC GTTTGGCAGTATTCATTGCC
B	*epaO-epaS*	TTATCTTCTGAAGAGAGCTG GTAAATCGAAACACTCTACG
L	*epaO-eprK*	ACGTTAGATTGCGCATTTCG AGTTAATGTTCGGTCTGAGG
C	*eivC-eivJ*	CCTTCATTTTCTTGCTCCTC TCTTGCGATATGGAGCATGC
D	*eivF-eivA*	GAAGATACTTCGAGGTCTCG AGGACAGAAGTTAACCCTAC

[a] PCR conditions were 25 cycles of 98°C for 1 min, 60°C for 10 min, and 72°C 150 min.

RESULTS

Structure of a new locus that encodes a second type III secretion system. Based on the DNA sequence of the EHEC O157 Sakai chromosome, we searched for a DNA sequence unique to EHEC O157 which was not present in the *E. coli* K-12 genome. One such sequence consisted of a 17-kb region inserted between the sequences homologous to those at min 64 and 65 of the *E. coli* K-12 chromosome (Fig. 1). The 17-kb insert showed no significant homology with the *E. coli* K-12 chromosome and contained 21 open reading frames (ORFs) (Table 4; Fig. 1). The G+C content of this region was 36.9%.

The deduced amino acid sequences of ORF2 to ORF5 and ORF7 to ORF20 were similar to those of proteins of the type III secretion system, which are required for the export of virulence determinants in a variety of bacterial pathogens of plants and animals. The predicted proteins of ORF2 to ORF5 are similar in organization and sequences to the products of the *prgH* to *prgK* genes of *Salmonella* SPI-1. The predicted proteins of ORF7 to ORF20 are similar in organization and sequences to the products of the *inv* and *spa* genes of *Salmonella* SPI-1 but not to those of *invB*. They are also homologous to the *mxi* genes of *Shigella flexneri* (data not shown). For example, the predicted amino acid sequence of ORF2 is 53% identical to that of PrgK of *S. enterica* serovar Typhimurium and 44% identical to that of MxiJ of *S. flexneri* 2a. The amino acid sequences of ORF3 to ORF5 are similar to those of PrgJ (31% identical), PrgI (64%), and PrgH (31%), respectively, of *S. enterica* serovar Typhimurium and to those of MxiI (32%), MxiH (58%), and MxiG (25%), respectively, of *S. flexneri* 2a. The entire organization and amino acid sequences of the ORFs in this 17-kb region are closely related to those of SPI-1 of *S. enterica* serovar Typhimurium (Fig. 1). In addition to these ORFs, the predicted protein of ORF1, which is located downstream of ORF2 and extends to the homologous sequence in *E. coli* K-12, showed significant similarity to the

TABLE 3. Primers used to detect pathogenic STEC genes[a]

Target gene	Sequence (5′ to 3′)	Cycle conditions (30 cycles)			Product size (bp)
		Denaturation	Annealing	Extension	
stx, common	GAGCGAAATAATTTATATGTG TGATGATGGCAATTCAGTAT	94°C, 1 min	58°C, 1 min	72°C, 1 min	518
stx_1	GCAGTTCGTGGCAAGAGCG GCGTCGCCAGCGCACTTG	94°C, 1 min	62°C, 1 min	72°C, 1 min	522
stx_2	AATTTATATGTGGCAGGGTTC CTTCACTGTAAATGTGTCATC	94°C, 1 min	52°C, 1 min	72°C, 1 min	806
stx_2 B subunit	GTTATACTGAATTGCCATCATC GTCATTATTAATCTGGAC	94°C, 1 min	52°C, 1 min	72°C, 1 min	435
eae, common	CCCGAATTCGGCACAAGCATAAGC CCCGGATCCGTCTCGCCAGTATTCG	94°C, 30 s	52°C, 1 min	72°C, 2 min	881
eae-α	CCCGAATTCGGCACAAGCATAAGC CCCGAATTCTTATTTTACACAAGTGGC	94°C, 45 s	48°C, 1 min	72°C, 2 min 30 s	2,807
eae-β	CCCGAATTCGGCACAAGCATAAGC CCCGTGATACCAGTACCAATTACGGTC	94°C, 45 s	48°C, 1 min	72°C, 2 min 30 s	2,287
eae-γ	CCCGAATTCGGCACAAGCATAAGC CCCGAATTCTTATTCTACACAAACCGC	94°C, 1 min	52°C, 1 min	72°C, 1 min	2,792
eae-ε	CCCGAATTCGGCACAAGCATAAGC AGCTCACTCGTAGATGACGGCAAGCG	94°C, 45 min	48°C, 1 min	72°C, 2 min 30 s	2,608
hlyA	GGTGCAGCAGAAAAAGTTGTAG TCTCGCCTGATAGTGTTTGGTA	94°C, 1 s	54°C, 1 min	72°C, 1 min 30 s	1,551

[a] Each primer set was described in other reports (1, 12, 13, 21, 25). *stx*, gene for Shiga toxin; stx_1, gene for Shiga toxin type 1; stx_2, gene for Shiga toxin type2; *eae*, gene for intimin; *hlyA*, gene for a plasmid-coded enterohemolysin.

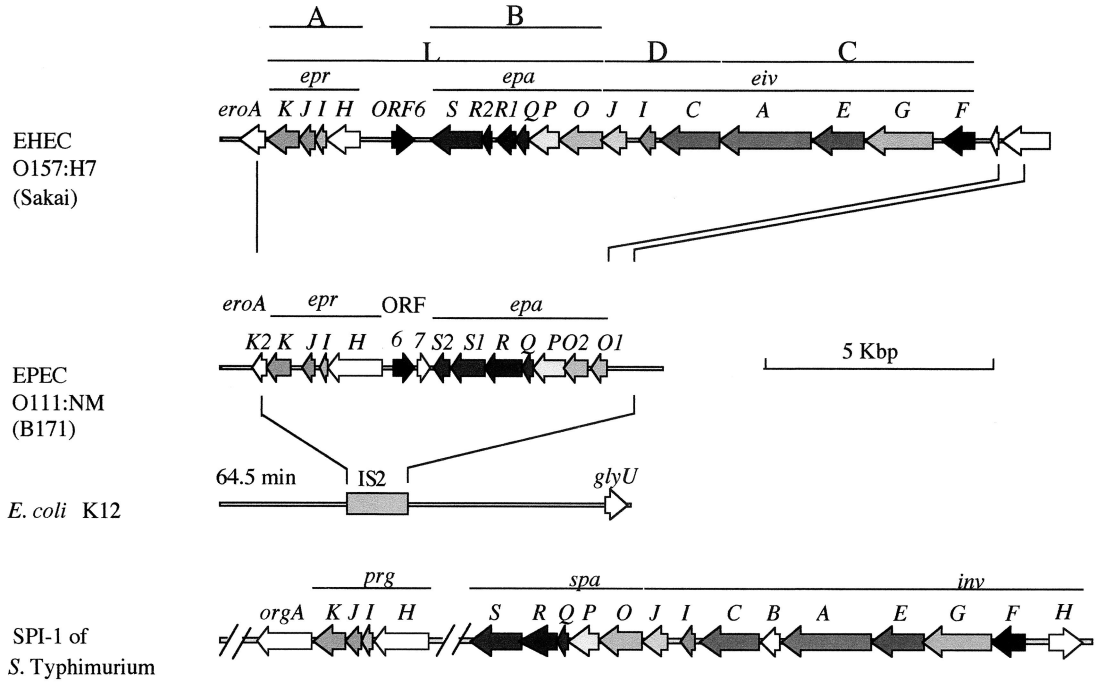

FIG. 1. Primary structure of ETT2 locus and homology with SPI-1 of *S. enterica* serovar Typhimurium. The locations of ORFs predicted from the nucleotide sequence are shown. The structures are of the following strains, from top to bottom: ETT2 of EHEC O157 Sakai, ETT2 of EPEC O111:NM B171, chromosome of *E. coli* K-12 W3110 at min 64.5, and SPI-1 of *S. enterica* serovar Typhimurium.

N-terminal part of the protein product of *orgA*, downstream of *prgK* in SPI-1 of *S. enterica* serovar Typhimurium. ORF8 and ORF9 showed significant similarity to SpaR of SPI-1, but the amino acid sequence of ORF9 was separated from ORF8, similar to the sequence in the C-terminal part of SpaR, by frame shifting in the ORF9-coding sequence. ORF5 was similar to PrgH in SPI-1 of *S. enterica* serovar Typhimurium, but the amino acid sequence corresponding to the N-terminal part of PrgH was lost from ORF5 because of the absence of the initiation codon, even though the DNA sequence upstream of

TABLE 4. ORFs in the ETT2 locus of EHEC O157 Sakai

ORF	Gene	SPI-1 of *S. enterica* serovar Typhimurium			No. of amino acids
		Homologous protein	% Identity/% similarity compared with EHEC O157 Sakai[a]	GenBank accession no.	
ORF1	*eorA*	OrgA	30/46 (179)	L33855	193
ORF2	*eprK*	PrgK	53/69 (232)	U21676	244
ORF3	*eprJ*	PrgJ	31/56 (87)	U21676	110
ORF4	*eprI*	PrgI	64/78 (76)	U21676	79
ORF5	*eprH*	PrgH	31/51 (225)	U21676	244
ORF6		NtrC	25/40 (107)	AF082873	166
ORF7	*epaS*	SpaS	54/72 (349)	X73525	373
ORF8	*epaR2*	SpaR	56/77 (65)	X73525	78
ORF9	*epaR1*	SpaR	44/65 (146)	X73525	150
ORF10	*epaQ*	SpaQ	68/84 (86)	U29364	86
ORF11	*epaP*	SpaP	54/62 (223)	U29364	221
ORF12	*epaO*	SpaO	32/48 (299)	U29350	318
ORF13	*eivJ*	InvJ	43/58 (72)	X73525	189
ORF14	*eivI*	InvI	31/55 (101)	U43304	111
ORF15	*eivC*	InvC	64/75 (372)	U08279	439
ORF16	*eivA*	InvA	58/75 (666)	M90846	666
ORF17	*eivE*	InvE	37/55 (323)	U43269	381
ORF18	*eivG*	InvG	53/68 (499)	X75302	505
ORF19	*eivF*	InvF	40/62 (203)	U08280	249
ORF20					59
ORF21					352

[a] The numbers in parentheses are the numbers of amino acids in each ORF.

TABLE 5. Conservation of ETT2 in EHEC and other pathogenic *E. coli* isolates

Strain	Result with probe DNA			Result by PCR			No. of isolates positive (total no. tested)
	A	B	C	L	D	C	
EHEC O157:H7	+	+	+	+	+	+	6 (6)[a]
EHEC O157:H−	+	+	+	+	+	+	1 (1)
EHEC O26	+	+	−	+	−	−	2 (2)[b]
EPEC	+	+	−	+	−	−	4 (4)
HB101	−	−	−	−	−	−	0 (1)

[a] Includes strain T-1 from New Caledonia, strains T-46 and K-159, from Canada, strain K-262 from Saitama, Japan, and strain. Obi-1 from Obihiro, Japan.
[b] Includes human strain AS1 and bovine strain H5-26.

ORF5 showed a high degree of similarity to that of the sequence upstream of the *prgH* gene in *S. enterica* serovar Typhimurium. Because of the high degree of similarity of each ORF to the genes in SPI-1 of *Salmonella* spp., the region from ORF5 to ORF2 was designated *eprH* to *eprK*, the region from ORF12 to ORF7 was designated *epaO* to *epaS*, and the region from ORF19 to ORF13 was designated *eivF* to *eivJ* (Fig. 1).

In SPI-1 of *S. enterica* serovar Typhimurium, the DNA between *spaS* and *prgH* encodes secreted proteins, chaperones, and transcriptional regulatory proteins; but the region between *epaS* and *eprH* of EHEC O157 Sakai contains no ORF similar to these genes. Instead, ORF6 had a low, but significant degree of similarity to several transcriptional regulatory proteins, including NtrC of *Herbaspirillum seropedicae* (25% identity and 40% similarity in a 107-amino-acid overlap) and UhpA of *S. enterica* serovar Typhimurium (29% identity and 49% similarity in a 68-amino-acid overlap). Consequently, the 17-kb region encodes the complete set of proteins for the type III secretion system apparatus but does not include the target or chaperone proteins. This locus contains a set of genes for the type III secretion system. Therefore, to distinguish this locus from a set of genes for type III secretion in LEE, we designated this locus ETT2, for *E. coli* type III secretion locus 2.

Conservation of ETT2 sequences among STEC strains. To examine the degree of conservation of the ETT2 sequence among pathogenic *E. coli* isolates from humans, Southern hybridization was performed with probes derived from the DNA corresponding to *eprH* to *eprK* (probe A in Fig. 1; Table 2), *epaO* to *epaS* (probe B), and *eivF* to *eivA* (probe C). All seven EHEC O157 strains had DNA sequences homologous to the sequence of the ETT2 locus (Table 5). However, EHEC O26 strains from humans contained DNA sequences homologous to probes A and B, but not to probe C (Table 5). The same pattern of hybridization observed with O26 was observed with four EPEC strains, but the DNA sequence that hybridized with ETT2-specific probes was not found in nonpathogenic strain *E. coli* HB101. To survey the distribution of the ETT2 locus among various STEC isolates, a PCR was performed by using three primer sets specific for regions L, C, and D in the ETT2 locus (Fig. 1; Table 2). It was found that the results of the PCR coincided with those of Southern hybridization (Tables 1 and 5).

Thus, a PCR survey was undertaken to determine the distribution of the ETT2 locus among 56 STEC isolates of various origins (Table 1). Sixteen STEC isolates had a complete ETT2 locus of L, D, and C regions (Table 1); 15 isolates were serotype O157, and one isolate, from a deer, was Out:H25 (where Out represents all O antiserum negative). Although four EPEC isolates and two O26 isolates had only an L region, according to PCR (Table 5), eight other STEC isolates were also of this type (Table 1): four were O26 and four were O119 ONT:H− (where ONT means some O antiserum positive but not identified), O153:H−, and Out:H−. Eight STEC strains, including five O157 strains, had a complete L region and either a C region or a D region. However, among these eight strains, STEC HI-7 and HI-N had complete L and C regions and a D region with a short insertion (Table 1). Six other STEC strains had complete L and D regions but no C region (Table 1). The ETT2 locus was identified in O157 Sakai, and all O157 isolates of human origin had a complete ETT2 locus. Each of the O157 strains originating from a housefly, salmon roe, and six cattle also had a complete ETT2 locus (Table 1; Fig. 2). However,

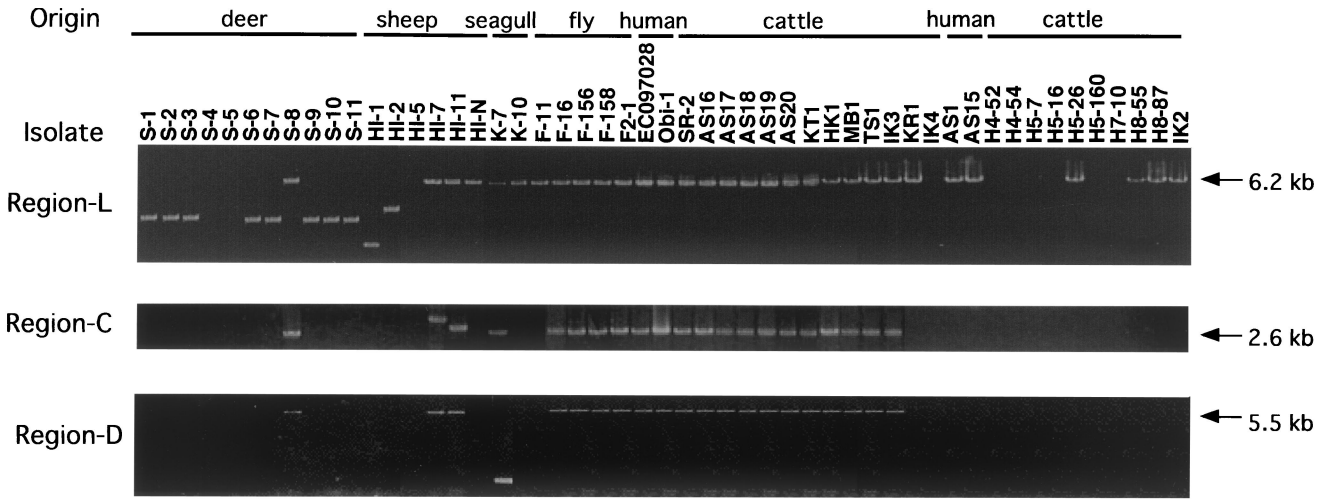

FIG. 2. Detection of regions L, C, and D among various STEC isolates by PCR. The arrows show the expected sizes of the fragments amplified from the DNA sequence.

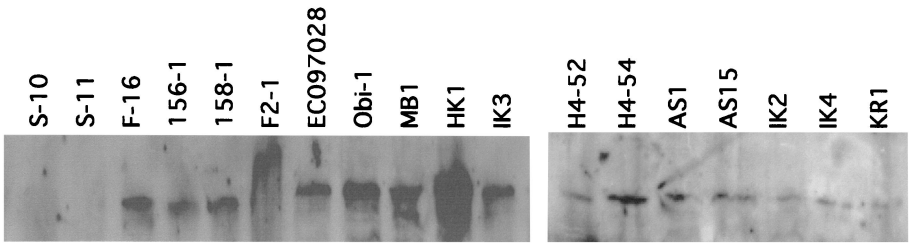

FIG. 3. Western blotting analysis of intimin proteins from different STEC isolates. Western blotting was performed by using antiserum against intimin encoded by *eae*-α. The results for representative STEC isolates are shown.

other O157 STEC strains from cattle had only the L and D regions (Table 1).

Distributions of other virulence factors among STEC strains. Previously, we examined the distributions of the *eae* and *hlyA* genes among STEC isolates of various origins using *eae* common primers and *hlyA*-specific primers (1, 2, 11). In the present study, the *eae* gene was classified by PCR and production of intimin was determined by Western blotting analysis with anti-intimin antibody. O157 STEC isolates commonly had *eae*-γ and produced intimin, regardless of their origins (Table 1; Fig. 3). In addition, all O157 STEC isolates except MB1, of cattle origin, had the *hlyA* gene. Among the other serotypes tested, O119 and O26 isolates commonly had the *eae*-β gene and produced intimin (Table 1; Fig. 3). Among them, eight isolates had the *hlyA* gene; among them, five isolates had the complete L region, *eae*, and *hlyA*. Most other STEC isolates had no *eae* gene and did not produce intimin (Table 1). STEC strains isolated from seagulls, deer, and sheep belonged to this group (Table 1).

DISCUSSION

A secondary type III secretion system, which was designated ETT2 (*E. coli* type III secretion system 2), was found in EHEC serotype O157:H7 isolates, in addition to the first system encoded by the LEE locus. This ETT2 locus contains 19 ORFs highly homologous with the SPI-1 type III secretion system of *S. enterica* serovar Typhimurium; the organization of those ORFs was also similar to that of SPI-1. However, slight differences from SPI-1 were found: (i) SpaR was separated into two ORFs, *epaR1* and *epaR2*, in the ETT2 locus, and (ii) InvB and InvH were absent from the ETT2 locus (Table 1). In the *Salmonella* SPI-1 system, InvH is necessary for localization of InvG onto the outer membrane, but it is not a component of the type III secretion apparatus (4), while the function of InvB remains unknown (6). Therefore, the ETT2 locus seems to encode a complete set of the type III secretion machinery. Indeed, the complete cloned sequence of ETT2 confers upon *E. coli* K-12 strain DH5-α the ability to secrete the EspB protein (T. Tobe, unpublished data).

In SPI-1, genes coding secreted (effector) proteins such as *sipBCDA*, the chaperon *sicP*, and the regulator *hilA* were found in the region between the *spa* and *prg* operons. Instead of these genes, ORF6, which had 40% similarity to the NtrC transcriptional regulator (23) of *H. seropedicae*, was found in ETT2 (Fig. 1), consistent with the ORF6 product also being a transcriptional regulator. Genes encoding target secreted proteins of the ETT2 locus may exist elsewhere on the O157 chromosome, but we could not identify them. However, a search for other sequences unique to O157 would provide the information needed to identify those genes.

In EPEC and O26 strains, the ETT2 locus was found on the chromosome, but the chromosomes of EHEC O157 strains contained only half of the ETT2 locus. Therefore, the ETT2 DNA segment of EPEC B171-8 (O111:NM) was isolated and its DNA sequence was determined (GenBank accession no. AB052736); it was also inserted at min 64.5 of the *E. coli* K-12 chromosome and was highly homologous with the L region of the ETT2 sequence in EHEC O157 Sakai (Fig. 1). The sequence corresponding to the *eiv* genes was absent from the ETT2 loci in EPEC and EHEC O26 strains (Fig. 1). All human STEC isolates had a complete or partial ETT2 locus, and notably, all O157 isolates had a complete ETT2 locus. Interestingly, those STEC and EPEC strains commonly had the L region with the *epr* and *epa* genes, suggesting that the ETT2 locus, especially the L-region locus, might be essential for the pathogenicity of O157 STEC isolates, as might the other known virulence factors encoded by the *eae* and *hlyA* genes. Although the association of the ETT2 locus with virulence in STEC was not clear, the pathogenicity of STEC might be estimated by PCR with the primer sets described in this study. For example, although non-O157 STEC strain S-8 had the complete ETT2 locus but not *stx*, *eae*, and *hlyA*, it might be pathogenic for humans. As O26 STEC strains AS1 and AS15 were isolated from humans, only the L region might be essential for pathogenicity. In this sense, strains KR1, H8-55, and IK2 from cattle would be virulent strains. Also, O157 strains lacking only the C region from five cattle also would be virulent for humans. As an *hlyA* gene associated with pathogenicity was encoded by a large plasmid, this gene would easily be lost. Therefore, as the *hlyA* gene would not be a good marker for the estimation of virulence (21), O157 strains MB1, H5-26, and H8-87 might have originally been virulent. Finally, cattle should be one of the main sources of infection for the transmission of STEC to humans. In this study, O157 isolates from a fly and salmon roe were closely similar to O157 isolates of human origin. However, as the salmon roe was the food that caused the O157 outbreak in Japan (15) and the fly was caught on cow manure on a cattle farm (11), these results were not unexpected.

The G+C content of the ETT2 sequence was very low (36.9%) compared with that of the *E. coli* chromosome (50.8%) (3), suggesting that the ETT2 locus has been acquired laterally. Furthermore, the high degree of similarity of ETT2 to

SPI-1 of *S. enterica* serovar Typhimurium suggests that ETT2 and SPI-1 may have originated from the same organism. However, the biological function of the ETT2 locus in the pathogenicity of STEC is not clear. The ETT2 locus might participate in the pathogenicity of these bacteria in a manner other than adherence to intestinal epithelial cells. Finally, as the ETT2 locus is common in human STEC isolates, it would be a good marker for molecular epidemiological surveys of STEC strains pathogenic for humans.

ACKNOWLEDGMENTS

We thank Ben Adler for critical suggestions on the manuscript.

This work was supported in part by grants from the Ministry of Health, Labour and Welfare (Research on Emerging and Re-Emerging Infectious Diseases), by a Grant-in Aid for Scientific Research from the Japanese Society for the Promotion of Science (grant 12575029), and by a grant from The 21st Century COE Program (A-1), Ministry of Education, Culture, Sports, Science and Technology of Japan.

REFERENCES

1. **Asakura, H., S. Makino, T. Shirahata, T. Tsukamoto, H. Kurazono, T. Ikeda, and K. Takesi.** 1998. Detection and genetical characterization of Shiga toxin-producing *Escherichia coli* from wild deer. Microbiol. Immunol. **42:** 815–822.
2. **Asakura, H., S. Makino, H. Kobori, M. Watarai, T. Shirahata, T. Ikeda, and K. Takeshi.** 2001. Phylogenetic diversity and similarity of active sites of Shiga toxin (Stx) in Shiga toxin-producing *Escherichia coli* (STEC) isolates from humans and animals. Epidemiol. Infect. **127:**27–36.
3. **Blattner, F. R., G. Plunkett III, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao.** 1997. The complete genome sequence of *Escherichia coli* K-12. Science **277:**1453–1474.
4. **Daefler, S., and M. Russel.** 1998. The *Salmonella typhimurium* InvH protein is an outer membrane lipoprotein required for the proper localization of InvG. Mol. Microbiol. **28:**1367–1380.
5. **Donnenberg, M. S., J. B. Kaper, and B. B. Finlay.** 1997. Interactions between enteropathogenic Escherichia coli and host epithelial cells. Trends Microbiol. **5:**109–114.
6. **Eichelberg, K., C. C. Ginocchio, and J. F. Galan.** 1994. Molecular and functional characterization of the *Salmonella typhimurium* invasion genes *invB* and *invC*: homology of InvC to the F0F1 ATPase family of proteins. J. Bacteriol. **176:**4501–4510.
7. **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.
8. **Galan, J. E., and R. Curtiss III.** 1989. Cloning and molecular characterization of genes whose products allow Salmonella typhimurium to penetrate tissue culture cells. Proc. Natl. Acad. Sci. USA **86:**6383–6387.
9. **Hayashi, T., K. Makino, M. Ohnishi, K. Kurokawa, K. Ishii, K. Yokoyama, C. G. Han, E. Ohtsubo, K. Nakayama, T. Murata, M. Tanaka, T. Tobe, T. Iida, H. Takami, T. Honda, C. Sasakawa, N. Ogasawara, T. Yasunaga, S. Kuhara, T. Shiba, M. Hattori, and H. Shinagawa.** 2001. Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12. DNA Res. **8:**11–22.
10. **Hueck, C. J.** 1998. Type III protein secretion systems in bacterial pathogens of animals and plants. Microbiol. Mol. Biol. Rev. **62:**379–433.
11. **Iwasa, M., S.-I. Makino, H. Asakura, H. Kobori, and Y. Morimoto.** 1999. Detection of *Escherichia coli* O157:H7 from *Musca domestica* (Diptera: Muscidae) at a cattle farm in Japan. J. Med. Entomol. **36:**108–112.
12. **Karch, H., H. Bohm, H. Schmidt, F. Gunzer, S. Aleksic, and J. Heesemann.** 1993. Clonal structure and pathogenicity of Shiga-like toxin-producing, sorbitol-fermenting *Escherichia coli* O157:H−. J. Clin. Microbiol. **31:**1200–1205.
13. **Levine, M. M., J. G. Xu, J. B. Kaper, H. Lior, V. Prado, R. Tall, J. Nataro, H. Karch, and I. K. Wachsmuth.** 1987. A DNA probe to identify enterohemorrhagic *Escherichia coli* of O157:H7 and other serotypes that cause hemorrhagic solitis and hemolytic uremic syndrome. J. Infect. Dis. **156:**176–182.
14. **Makino, K., K. Ishii, T. Yasunaga, M. Hattori, K. Yokoyama, C. H. Yatsudo, Y. Kubota, Y. Yamaichi, T. Iida, K. Yamamoto, T. Honda, C. G. Han, E. Ohtsubo, M. Kasamatsu, T. Hayashi, S. Kuhara, and H. Shinagawa.** 1998. Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrhagic *Escherichia coli* O157:H7 derived from Sakai outbreak. DNA Res. **5:**1–9.
15. **Makino, S.-I., T. Kii, H. Asakura, T. Shirahata, T. Ikeda, K. Takeshi, and K. Itho.** 2000. Does enterohemorrhagic *Escherichia coli* O157:H7 enter the viable but nonculturable (VNC) state in salted salmon roe? Appl. Environ. Microbiol. **66:**5536–5539.
16. **Makino, S.-I., K. Yamamoto, S. Murakami, T. Shirahata, K. Uemera, T. Sawada, H. Wakamoto, and Y. Morita.** 1998. Properties of repeat domain found in a novel protective antigen, SpaA, of *Erysipelothrix rhusiopathiae*. Microb. Pathog. **723:**101–109.
17. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
18. **McDaniel, T. K., and J. B. Kaper.** 1997. A cloned pathogenicity island from enteropathogenic *Escherichia coli* confers the attaching and effacing phenotype on *E. coli* K-12. Mol. Microbiol. **23:**399–407.
19. **Nataro, J. P., and J. B. Kaper.** 1998. Diarrheagenic *Escherichia coli*. Clin. Microbiol. Rev. **11:**142–201.
20. **Puente, J. L., D. Bieber, S. W. Ramer, W. Murray, and G. K. Schoolnik.** 1996. The bundle-forming pili of enteropathogenic *Escherichia coli*: transcriptional regulation by environmental signals. Mol. Microbiol. **20:**87–100.
21. **Schmidt, H., H. Russmann, and H. Karch.** 1993. Virulence determinants in nontoxinogenic *Escherichia coli* O157 strains that cause infantile diarrhea. Infect. Immun. **61:**4894–4898.
22. **Shea, J. E., M. Hensel, C. Gleeson, and D. W. Holden.** 1996. Identification of a virulence locus encoding a second type III secretion system in *Salmonella typhimurium*. Proc. Natl. Acad. Sci. USA **93:**2593–2597.
23. **Souza, E. M., F. O. Pedrosa, L. U. Rigo, H. B. Machado, and M. G. Yates.** 2000. Expression of the nifA gene of *Herbaspirillum seropedicae*: role of the NtrC and NifA binding sites and of the −24/−12 promoter element. Microbiology **146:**1407–1418.
24. **Woestyn, S., A. Allaoui, P. Wattiau, and G. R. Cornelis.** 1994. YscU, a *Yersinia enterocolitica* inner membrane protein involved in Yop secretion. J. Bacteriol. **176:**1561–1569.
25. **Yamasaki, S., Z. Lin, and H. Shirai.** 1996. Typing of verotoxins by DNA colony hybridization with poly- and oligonucleotide probes, a bead-enzyme-linked immunosorbent assay, and polymerase chain reaction. Microbiol. Immunol. **40:**345–352.