AEM
Journals.ASM.org

Distribution of Pathogenicity Islands OI-122, OI-43/48, and OI-57 and a High-Pathogenicity Island in Shiga Toxin-Producing *Escherichia coli*

Wenting Ju,[a] Jinling Shen,[a,c] Magaly Toro,[a] Shaohua Zhao,[d] Jianghong Meng[a,b]

Department of Nutrition and Food Science, University of Maryland, College Park, Maryland, USA[a]; Joint Institute of Food Safety and Applied Nutrition, University of Maryland, College Park, Maryland, USA[b]; College of Food Science and Engineering, Northwest A&F University, Yangling, Shaanxi Province, China[c]; Division of Animal and Food Microbiology, Office of Research, Center for Veterinary Medicine, Food and Drug Administration, Laurel, Maryland, USA[d]

Pathogenicity islands (PAIs) play an important role in Shiga toxin-producing *Escherichia coli* (STEC) pathogenicity. The distribution of PAIs OI-122, OI-43/48, and OI-57 and a high-pathogenicity island (HPI) were determined among 98 STEC strains assigned to seropathotypes (SPTs) A to E. PCR and PCR-restriction fragment length polymorphism assays were used to identify 14 virulence genes that belonged to the four PAIs and to subtype *eae* and *stx* genes, respectively. Phylogenetic trees were constructed based on the sequences of *pagC* among 34 STEC strains and *iha* among 67 diverse pathogenic *E. coli*, respectively. Statistical analysis demonstrated that the prevalences of OI-122 (55.82%) and OI-57 (82.35%) were significantly greater in SPTs (i.e., SPTs A, B, and C) that are frequently associated with severe disease than in other SPTs. *terC* (62.5%) and *ureC* (62.5%) in OI-43/48 were also significantly more prevalent in SPTs A, B, and C than in SPTs D and E. In addition, OI-122, OI-57, and OI-43/48 and their associated virulence genes (except *iha*) were found to be primarily associated with *eae*-positive STEC, whereas HPI occurred independently of the *eae* presence. The strong association of OI-122, OI-43/48, and OI-57 with *eae*-positive STEC suggests in part that different pathogenic mechanisms exist between *eae*-positive and *eae*-negative STEC strains. Virulence genes in PAIs that are associated with severe diseases can be used as potential markers to aid in identifying highly virulent STEC.

Shiga toxin-producing *Escherichia coli* (STEC) can cause human illnesses ranging from self-limiting diarrhea to life-threatening diseases such as hemolytic-uremic syndrome (HUS), a leading cause of kidney failure in children (1). *E. coli* O157:H7 is the single serotype that causes most STEC outbreaks and HUS cases. Like O157, non-O157 STEC can also cause severe diseases and food-borne outbreaks (1). More than 470 non-O157 STEC serotypes have been associated with human illness (2), and public health concerns regarding non-O157 STEC are increasing (1). Estimations indicate that non-O157 STEC strains cause 112,752 illness each year in the United States, almost twice the number of O157:H7 illnesses (63,153) (3). Although some non-O157 STEC strains have been associated with disease symptoms indistinguishable from O157:H7, not all STEC serotypes can cause HUS and outbreaks, and some STEC serotypes have never been reported to be related to any human illness (4). The scientific basis for this difference, however, is poorly understood.

Increasing evidence shows that differences in virulence between pathogenic and nonpathogenic bacterial strains can be attributed in part to virulence genes located in pathogenicity islands (PAIs) (5). PAIs usually contain blocks of virulence genes and are >10 kb (6). Several PAIs have been identified and characterized in STEC. A chromosomal pathogenicity island termed the locus of enterocyte effacement (LEE) was identified in *E. coli* O157:H7 strain EDL933, which encodes a type III secretion system, as well as virulence genes (*eae* and *tir*) associated with the intimate attachment of bacteria to intestinal epithelial cells (4). LEE appears to confer enhanced virulence, since LEE-positive STEC are much more commonly associated with HUS and outbreaks than LEE-negative STEC (5). However, some LEE-positive STEC serotypes have never been associated with disease, and some LEE-negative STEC can cause HUS and outbreaks, indicating that virulence factors other than those in LEE may contribute to pathogenesis of STEC (5).

Pathogenicity island OI-122 is also well characterized in O157:H7 (5, 7). OI-122 is a 23-kb PAI consisting of three modules (5, 8, 9). Z4321 is located in module 1 and encodes a protein sharing 46% similarity with the *phoP*-activated gene C (*pagC*) of *Salmonella enterica* serovar Typhimurium (5, 9). Z4326, Z4328, and Z4329 are located in module 2. Z4326 (*sen*) encodes a protein that shares 38.2% similarity to *Shigella flexneri* enterotoxin 2 (5), whereas Z4328 and Z4329 encode proteins that have 89 and 86% similarity to non-LEE-encoded effectors NleB and NleE, respectively (9). The enterohemorrhagic *E. coli* factor for adherence (Efa), which is involved in epithelial cell adhesion and inhibiting the proliferation of bovine peripheral blood lymphocytes, is located in module 3 (10).

OI-43 and OI-48 are duplicate genomic islands found in EDL933 (8). OI-43/48 genes are divided into three functional groups: a seven-gene cluster *ureDABCEFG* that encodes urease and accessory proteins hydrolyzing urea to ammonia and carbon dioxide; telluride resistance genes *terZABCDEF* (11); and two putative adhesion genes, *iha* (iron-regulated gene A) and *aidA-1* (autotransporter adhesin involved in diffuse adherence) (12).

Received 13 January 2013 Accepted 19 March 2013
Published ahead of print 22 March 2013
Address correspondence to Jianghong Meng, jmeng@umd.edu.
Supplemental material for this article may be found at http://dx.doi.org/10.1128/AEM.03661-12.

doi:10.1128/AEM.03661-12

In EDL933, OI-57 contains non-LEE-encoded effector genes *nleG2-3*, *nleG6-2*, and *nleG5-2* (13, 14). NleG proteins are E3 ubiquitin ligases analogous to RING finger and U-box enzymes in eukaryotes. Although the exact functions of NleG2-3, NleG6-2, and NleG 5-2 are still unclear, similar proteins have been identified as effectors that suppress immune response from the host (15).

The high-pathogenicity island (HPI) was first detected in *Yersinia pestis* and other highly virulent *Yersinia* species and encodes a siderophore (yersiniabactin)-mediated iron-uptake system (16). HPI is required for full virulence expression in *Yersinia* (16) and contains two main virulence genes, *fyuA* and *irp2*. FyuA is an outer membrane protein acting as a receptor for ferric-yersiniabactin uptake and for bacteriocin pesticin, whereas Irp2 is involved in yersiniabactin synthesis (17). An orthologous and highly conserved HPI is widely distributed among different species and genera of the family *Enterobacteriaceae* (16).

Few studies to date have investigated PAIs other than LEE in STEC. Since PAIs are normally absent in nonpathogenic strains of the same or closely related species, they may serve as useful markers to distinguish highly virulent strains from less-virulent or harmless strains (5, 6). In addition, PAIs can be used to identify new and emerging pathogenic bacteria. We report here the distribution of OI-122, OI-43/48, OI-57, and HPI and their virulence genes in STEC and evaluate the association of the PAIs and individual virulence genes with STEC seropathotypes (SPTs) linked to severe diseases and outbreaks. In addition, the association of the four PAIs with LEE was determined.

MATERIALS AND METHODS

Bacterial strains. A total of 98 STEC strains from humans, animals, and food were used in the present study (see Table S1 in the supplemental material). Strains were classified into SPTs A to E according to the criteria described by Karmali et al. (5). The assignment of SPTs was based on published references (5, 14, 18) and a large online database on non-O157 STEC (http://www.lugo.usc.es/ecoli/serotiposhum.htm).

***stx* and *eae* subtyping.** *stx* and *eae* subtypes were determined using PCR-restriction fragment length polymorphism (RFLP) analysis (19, 20), and stx_{2dact} was confirmed by PCR as previously described (21). Genomic DNA was extracted using boiling method as previously described (22, 23). STEC strains S1191 (stx_{2e}), EDL933 (stx_{1a} and stx_{2a}), E32511 (stx_{2c}), EH250 (stx_{2b}), B2F1 (stx_{2dact}), and N15018 (stx_{1c}) were used as positive controls for the *stx* subtyping; STEC strains 86-24 (gamma 1), EDL933 (gamma 1), TW06584 (kappa), E2348-69 (alpha), TW07920 (epsilon), RDEC-1 (beta), TW10366 (rho), TW03501 (iota), TW07892 (eta), and TW01387 (gamma 2/theta) were used as positive controls for the *eae* subtyping. *E. coli* K-12 was used as a negative control strain for both *stx* and *eae* subtyping.

Presence of OI-122, OI-43/48, O-57, and HPI. PCR assays were used to determine the presence of 14 virulence genes in STEC OI-122, OI-43/48, OI-57, and HPI as described previously (5, 11, 13, 24, 25). The presence of a PAI was determined by several marker genes located in different regions of the island, including *pagC*, *sen*, *nleB*, *efa-1*, and *efa-2* for OI-122; *terC*, *ureC*, *iha*, and *aidA-1* for OI-43/48; *nle2-3*, *nleG6-2*, and *nleG5-2* for OI-57; and *irp2* and *fyuA* for HPI. PCR was performed in a 25-μl reaction mixture, containing 2 μl of DNA template, 2.5 μl of 10× PCR buffer, 2 μl of a 25-mmol liter^{-1} solution of $MgCl_2$, 2 μl of a 1.25-mmol liter^{-1} deoxynucleoside triphosphate mixture, 0.125 μl of a 5-U μl^{-1} AmpliTaq Gold DNA polymerase mixture (Applied Biosystems, Branchburg, NJ), and 0.2 μl of a 50-pmol μl^{-1} concentration of each primer. *E. coli* O157:H7 EDL 933 was used as a positive control for the virulence genes of OI-122, OI-43/48, and OI-57 and *E. coli* O26:H11 SJ-13 for the virulence genes of HPI. *E. coli* K-12 was used as a negative control for all PCR assays.

Phylogenetic and sequence analysis. *iha* and *pagC* were the only two genes that were highly prevalent in both *eae*-positive and *eae*-negative STEC strains. To determine the evolutionary relationship between the two groups of STEC, *iha* and *pagC* were selected for phylogenetic analysis studies. *iha* sequences from 67 *E. coli* and *Shigella* strains were obtained from GenBank. A multiple sequence alignment of *iha* was performed using CLUSTAL W in MEGA 5.05, and a maximum-likelihood phylogenetic tree was generated using the general time reversible model (26). A bootstrapping of 2,000 replicates was used to estimate the confidence of the branching patterns of the phylogenetic tree using *iha* of *E. coli* SMS-3-5 as the phylogenetic tree's root.

In addition, PCR was used to amplify *pagC* of OI-122 from 12 selected STEC strains representing different serotypes, as described by Konczy et al. (9). PCR products were sequenced by GeneWiz (Germantown, MD). Twenty-two *pagC* sequences representing different STEC serotypes and one *Citrobacter* strain were downloaded from GenBank. The *pagC* sequences were cropped to 446 bp prior to alignment. Phylogenetic analysis was performed using CLUSTAL W within MEGA 5.05 (26). A phylogenetic tree based on *pagC* sequences was constructed using maximum-likelihood methods by MEGA 5.05 with bootstrapping of 2,000 replicates using *pagC* of *Citrobacter rodentium* IC168 as the tree's root.

Statistical analysis. A chi-square or Fisher exact test was used for data analysis using SAS9.2 (SAS Institute, Cary, NC). A P value of <0.01 was considered statistically significant.

RESULTS

Distribution of OI-122, OI-43/48, OI-57, and HPI in STEC strains. The 98 STEC strains were classified into SPTs A to E (see Table S1 in the supplemental material). Overall, the prevalence of OI-122 and OI-57 decreased progressively from SPT A to SPT E (see Fig. S1 in the supplemental material). The prevalences of OI-122 and OI-57 were significantly greater in SPTs associated with severe diseases (SPTs A, B, and C) and outbreaks (SPTs A and B) than in other SPTs ($P < 0.0001$) (Table 1). Although the prevalence of OI-43/48 was greater in SPTs associated with HUS (SPTs A, B, and C) and outbreaks (SPTs A and B) than in other SPTs, the differences were not statistically significant ($P = 0.1356$ and 0.02, respectively). HPI was not found in SPT A (O157) but was almost evenly distributed from SPT B to SPT E (see Fig. S1 in the supplemental material).

sen, *nleB*, *efa-1*, *efa-2*, *terC*, *ureC*, *nleG2-1*, *nleG5-2*, and *nleG6-2* were significantly more prevalent in SPTs A and B than in SPTs C, D, and E (Table 1). *pagC*, *sen*, *nleB*, *efa-1*, *efa-2*, *terC*, *ureC*, *nleG2-1*, *nleG5-2*, and *nleG6-2* were statistically more prevalent in SPTs A, B, and C than in SPTs D and E (Table 1). Although *aidA-1* was more prevalent in SPTs A and B than in SPTs C, D, and E, the difference was not statistically significant ($P = 0.27$). *iha*, *fyuA*, and *irp2* were less prevalent in SPTs A, B, and C than in SPTs D and E, but the differences were not statistically significant.

Distribution of OI-122, OI-43/48, OI-57, and HPI in EHEC O157. PAIs showed three patterns of distribution in EHEC O157 (see Table S1 in the supplemental material). In the case of β-glucuronidase (GUD)-negative O157:H7, four strains all contained marker genes for OI-122, OI-57, and OI-43/48. In GUD-positive O157: H7, none of the five strains carried *efa-1* and *efa-2* (located at the third module of OI-122) or *aidA-1* (located at the end of OI-43/48). Sorbitol-fermenting O157:NM strains contained all of the virulence genes of OI-122 and OI-57 but were negative for all of the OI-43/48 virulence marker genes, indicating the absence of OI-43/48 in O157:NM.

TABLE 1 Association of PAIs and virulence genes with SPTs related to outbreak (SPTs A and B), severe disease (SPTs A, B, and C), and LEE

PAI	Gene	Prevalence (%)[a]					
		Association with SPTs related to outbreak		Association with SPTs related to severe disease		Association with LEE	
		In SPTs A and B ($n = 34$)	In SPTs C, D and E ($n = 64$)	In STEC SPTs A, B, and C ($n = 56$)	In STEC in SPTs D and E ($n = 42$)	In *eae*-positive STEC ($n = 54$)	In *eae*-negative STEC ($n = 44$)
OI-122		55.82*	17.18*	46.43†	9.52†	55.56‡	0‡
	pagC	70.59	46.88	69.64 †	35.71†	64.81‡	38.64‡
	sen	100.00*	31.25*	76.79 †	26.45†	100.00‡	2.27‡
	efa-1	82.35*	31.25*	66.07†	30.95†	88.89‡	6.82‡
	efa-2	82.35*	31.25*	66.07†	26.45†	88.89‡	2.27‡
	nleB	100.00*	31.25*	76.79†	26.45†	100.00‡	2.27‡
OI-43/48		32.35	12.50	26.79	14.28	37.03‡	2.27‡
	terC	76.47*	35.94*	62.50†	33.33†	81.48‡	11.36‡
	ureC	76.47*	31.25*	62.50†	26.45†	81.48‡	4.54‡
	iha	64.70	70.31	64.29	73.80	57.41‡	84.09‡
	aidA-1	38.23	32.81	39.28	28.57	65.91‡	18.18‡
OI-57		82.35*	21.86*	60.71†	19.05†	75.93‡	2.27‡
	nleG2-3	97.06*	31.25*	73.21†	28.57†	94.44‡	6.82‡
	nleG5-2	82.35*	21.88*	60.71†	19.04†	85.19‡	2.27‡
	nleG6-2	97.18*	28.13*	66.07†	28.57†	75.93‡	4.56‡
HPI		17.65	25.00	17.86	28.57	25.92	18.18
	fyuA	17.65	23.44	16.07	28.57	25.92	18.18
	irp2	17.65	23.44	16.07	28.57	25.92	18.18

[a] *, Statistically significant difference between SPTs A and B compared to SPTs C, D, and E; †, statistically significant difference between SPTs A, B, and C compared to SPTs D and E; ‡, statistically significant difference between *eae*-positive and *eae*-negative STEC. A *P* value of <0.01 was considered statistically significant.

In addition, none of the O157:H7 and O157:NM strains were positive for HPI virulence genes.

Association of OI-122, OI-43/48, OI-57, and HPI with *eae*. We compared the distribution of virulence genes of OI-122, OI-43/48, OI-57, and HPI between *eae*-positive and *eae*-negative STEC strains. All virulence genes of OI-122 and OI-57 (*pagC*, *sen*, *nleB*, *efa-1*, *efa-2*, *nleG2-3*, *nleG5-2*, and *nleG6-2*) were highly prevalent in *eae*-positive strains (Table 1 and see Table S1 in the supplemental material). However, these genes, with the exception of *pagC*, were less prevalent in *eae*-negative STEC (Table 1 and see Table S1 in the supplemental material). Although 38.6% of *eae*-negative STEC strains were positive for *pagC*, its prevalence was significantly higher (64.8%) in *eae*-positive STEC ($P = 0.005$). There was no apparent physical or functional relationship identified between OI-43/48 and LEE, but three OI-43/48 virulence genes (*ureC*, *terC*, and *aidA-1*) were mainly associated with the presence of *eae* ($P < 0.0001$). On the other hand, *iha* was more prevalent in *eae*-negative than in *eae*-positive STEC strains ($P = 0.007$). For HPI, there were no significant differences in the distribution of *fyuA* or *irp2* between *eae*-positive and *eae*-negative STEC ($P = 0.36$).

Phylogenetic analysis of *iha* from diverse pathogenic *E. coli*. A phylogenetic tree based on *iha* separated *eae*-positive and *eae*-negative STEC strains into two distinct clades (Fig. 1). In clade I, two subgroups—Ia and Ib—shared at least 98.0% sequence similarity. *eae*-positive EHEC serotypes highly associated with outbreaks and severe diseases were located in clade I (O157:H7, O26:H11, O103:H2, O111:NM, and O145:H28). These sequences shared at least 99% similarity and clustered together with *iha* from other pathogenic *E. coli*, including enteropathogenic *E. coli* (EPEC), enteroinvasive *E. coli* (EIEC), enteroaggregative *E. coli* (EAEC), enterotoxigenic *E. coli* (ETEC), uropathogenic *E. coli* (UPEC), neonatal meningitis *E. coli* (NMEC), and Shiga toxin-producing EAEC O104:H4 (from a German outbreak in 2011). *iha* sequences from the O26:H11, O111:H11, and O111:NM strains formed subgroup Ib and shared at least 98.0% sequence similarity with subgroup Ia. Interestingly, strains DEC10A (O26:H11), DEC10C (O26:H11), 11368 (O26:H11), and DEC8C (O111:NM) carried two *iha* that clustered separately in Ia and Ib.

All 15 *iha* sequences from *eae*-negative STEC clustered together to form clade II. Multiple sequence alignments demonstrated that *iha* from *eae*-negative STEC shared only 91.1–93.6% sequence similarity with *iha* from clade I. *iha* from subgroups IIa and IIb shared only 93.8 to 94.3% sequence similarity. As in some *eae*-positive strains, *eae*-negative STEC strains CL-3 (O113:H21), 96.0497 (O91:H21), and B2F1 (O91:H21) also carried two *iha* genes that clustered separately in subgroups IIa and IIb.

Phylogenetic and sequence analysis of *pagC*. The *pagC* phylogenetic tree showed four clades (Fig. 2). *eae*-positive *E. coli* STEC formed a single clade with EPEC and ETEC, whereas the *eae*-negative STEC strains formed two clades, along with one strain that clustered with a *C. rodentium* strain. We identified 15 single nucleotide polymorphisms and one indel among the 35 *pagC* sequences. Sequence analysis revealed that an insertion of adenine at nucleotide 388 in two O103:H25 strains, two O45:H2 strains, one O103:H2 strain, and one O103:H6 strain led to a frameshift mutation and that a premature stop codon truncated the protein at the third loop, resulting in the loss of the fourth and last loops.

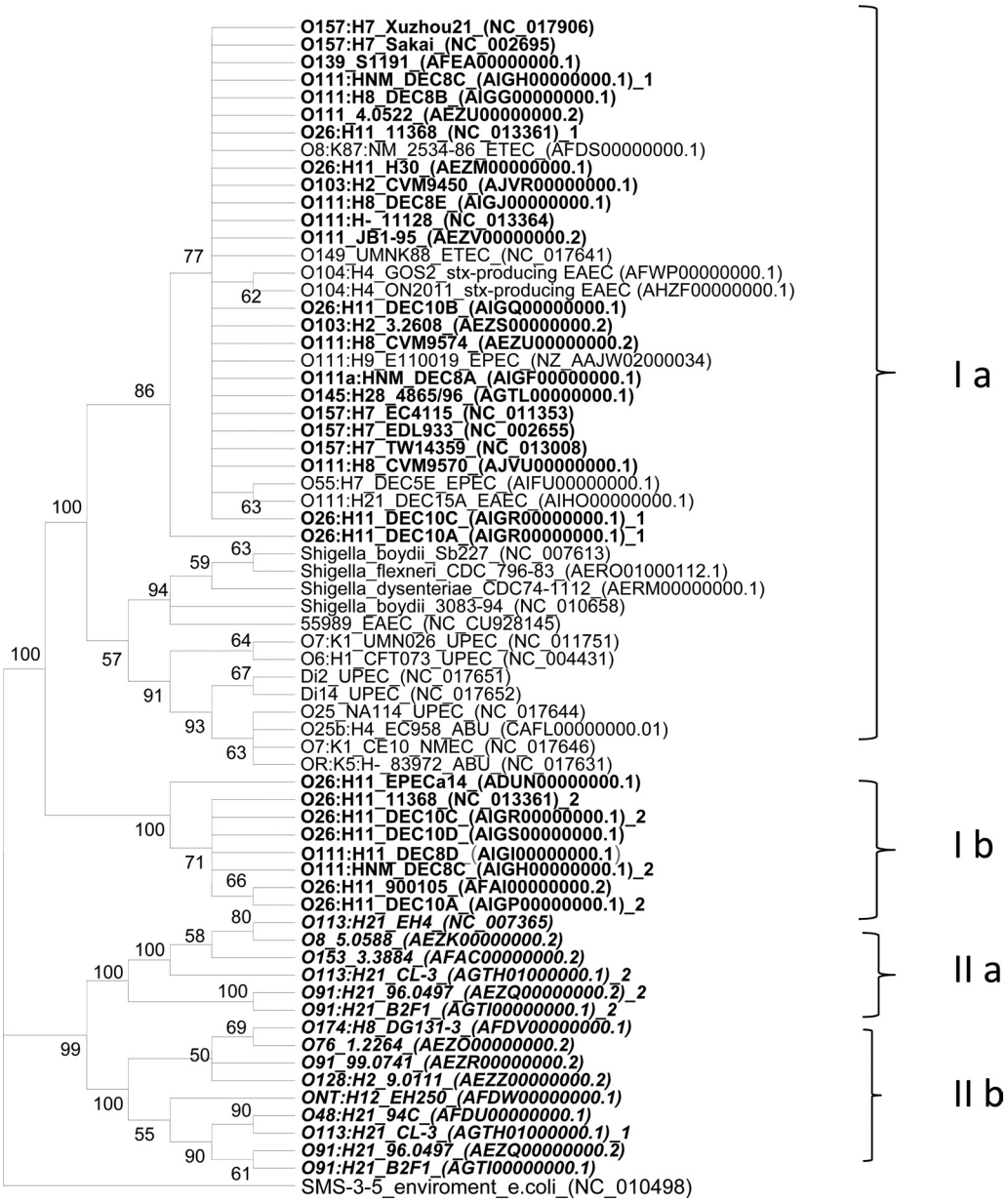

FIG 1 Phylogenetic tree based on *iha* sequences from 67 *E. coli* and *Shigella* strains. *iha* sequences were aligned, and a tree was constructed using the maximum-likelihood method with 2,000 iterations utilizing MEGA 5.05 (26). *iha* sequences from *eae*-positive and *eae*-negative STEC strains segregated into two distinct clades: clade I (with subgroup Ia and Ib) and clade II (with subgroups IIa and IIb). *iha* sequences from *eae*-negative STEC are marked in boldface italic type, and *eae*-positive STEC strains are marked in boldface regular type. EPEC, enteropathogenic *E. coli*; EIEC, enteroinvasive *E. coli*; EAEC, enteroaggregative *E. coli*; stx-producing EAEC, Shiga toxin-producing EAEC; ETEC, enterotoxigenic *E. coli*; UPEC, uropathogenic *E. coli*; NMEC, neonatal meningitis *E. coli*; ABU, asymptomatic bacteriuria *E. coli.*

DISCUSSION

In the present study, STEC PAIs OI-122 and OI-57 were found to be highly associated with seropathotypes that can cause severe disease and outbreaks, as previously demonstrated (7, 13, 14). Several OI-122 virulence factors play important roles in bacterial pathogenesis. For example, PagC can promote the survival of *Salmonella* within macrophages (5, 9). Efa is an adhesion protein originally described in some EHEC strains (27). The *efa-1* gene is almost identical to *lifA*, an EPEC gene encoding lymphostatin (LifA) (28), which inhibits the proliferation of mitogen-activated lymphocytes and the synthesis of proinflammatory cytokines (28). Efa1/LifA also contributes to EPEC adherence to epithelial cells and is critical for intestinal colonization by *C. rodentium* (29). NleB is required for full colonization and colonic hyperplasia in mice, and a mutation of *nleB* abolished the lethality of *C. rodentium* in C3H/HeJ mice (7, 30).

Whereas OI-122 is highly related to colonization and suppression of the host immune system, the function of OI-57 is largely unknown. Wu et al. (15) determined that NleG-like proteins and U-box enzymes in eukaryotes. Although the targets of the OI-57

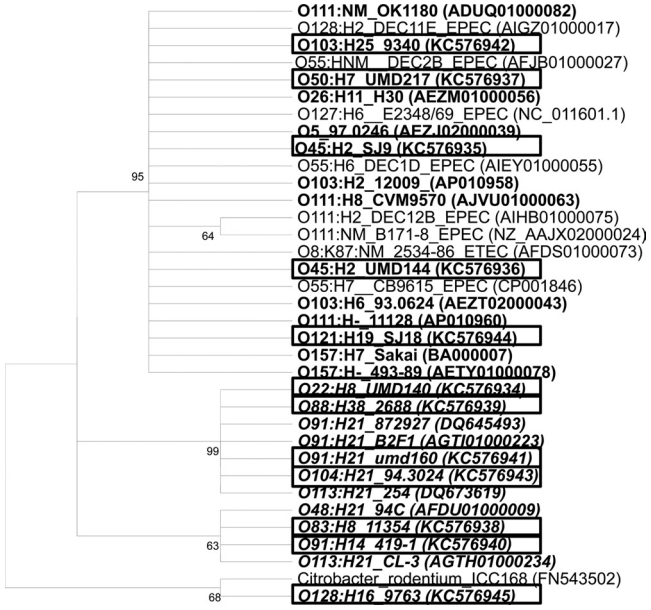

FIG 2 Phylogenetic tree based on *pagC* sequences from 34 pathogenic *E. coli* strains. *pagC* sequences were aligned, and a tree was constructed using the maximum-likelihood method with 2,000 iterations utilizing MEGA 5.05 (26). *pagC* sequences from *eae*-negative STEC are marked in boldface italic type, and *eae*-positive STEC strains were marked in boldface regular type. *pagC* genes sequenced in this study are marked by black frames. EPEC, enteropathogenic *E. coli*; ETEC, enterotoxigenic *E. coli*.

Nle effectors are unknown, several similar effectors are primarily involved in suppressing host immune response by degrading immune-related host proteins (15). Thus, it is possible that OI-57, similar to OI-122, would be also related to suppression of the host immune system.

In addition to the virulence genes in OI-122 and OI-57, the genes *ureC* and *terC*, located on OI-43/48, were also highly associated with seropathotypes related to severe disease and outbreaks. Urease has been confirmed as an important virulence factor in several bacterial species, such as *Helicobacter pylori*, *Yersinia enterocolitica*, *Proteus mirabilis*, *Brucella* species, and *Klebsiella pneumoniae* (31). Mutation of *ureC* has led to a reduced adherence of EHEC O157:H7 in ligated pig intestine (12). A recent study by Steyert and Kaper (32) revealed that strains with nonfunctional urease were 2-fold less likely to survive passage through the stomach and had a reduced ability to colonize the mouse intestinal tract compared to urease-positive strains. These data demonstrate that urease can help STEC strains survive in the stomach and enhance its competitiveness in colonization in calf and human intestinal tracts. The role of tellurite resistance genes (*terZABCDEF*) in STEC is still not well understood. Yin et al. (12) showed that mutation of the *ter* cluster in O157:H7 led to fewer adherence to epithelial cells and smaller bacterial clusters compared to wild-type strains. Therefore, *ter* genes might encode an adhesin or a gene product that promotes the function of adhesion(s). In addition, tellurite salts are strong oxidative agents, and it is possible that *ter* genes might offer a selective advantage in the host environment and aid STEC in general stress response (12).

Interestingly, *ureC* has been more frequently found in *eae*-positive STEC (113/132) than in *eae*-negative strains (4/70), although no physical linkage of *ureC* and *eae* has been identified (33). The prevalence of *ureC* in *eae*-positive STEC (45/55) was significantly higher than in *eae*-negative STEC (2/44) ($P < 0.0001$). Similarly, *terC* was also more prevalent in *eae*-positive STEC (45/55) than *eae*-negative strains (5/44) ($P < 0.0001$). Even though OI-43/48 and LEE are physically distant, our observations indicated that there might be a functional relationship between them.

The arrangement of OI-122 genes was found to be serotype dependent, and all O157:H7 strains have a complete OI-122 (5, 9). However, we found that two patterns of OI-122 existed in O157:H7. An incomplete OI-122 lacking the third module was identified in all GUD-positive O157:H7 strains. In addition, *aidA-1* of OI-43/48 was absent in GUD-positive O157:H7.

Most OI-122, OI-43/48, and OI-57 virulence genes (*pagC*, *sen*, *nleB*, *efa-1*, *efa-2*, *terC*, *ureC*, *iha*, *aidA-1*, *nleG2-3*, *nleG6-2*, and *nleG5-2*) were highly prevalent in *eae*-positive STEC. However, they were largely absent in *eae*-negative STEC, with the exception of *pagC* and *iha*. Phylogenetic analysis revealed that *iha* genes from *eae*-positive STEC had high similarity (99.6%), whereas they had lower sequence similarity (91.1 to 93.6%) to *iha genes* from *eae*-negative STEC, indicating that *iha* from *eae*-positive and *eae*-negative STEC strains may have evolved independently or have different origins. Such a difference also existed in *pagC* between *eae*-positive and *eae*-negative STEC strains. Schmidt et al. (34) reported that *iha* was carried by a 33,014-bp PAI in STEC serotype O91:H− strains (*eae* negative). In addition, *iha* was found in plasmid pO113 of STEC serotype O113:H21 (*eae* negative) (35). Moreover, Shen et al. (36, 37) reported that *pagC* was identified within a mosaic PAI from STEC O113:H21 strain CL-3 (*eae* negative). Thus, the higher prevalence of *iha* and *pagC* in the *eae*-negative STEC strains, compared to other virulence marker genes in the present study, is likely due to the presence of the same or similar PAIs and/or plasmids, as previously described. The similar prevalence of *iha* genes in the seropathotypes highly associated with severe diseases and other seropathotypes indicates that *iha* is not related to severe clinical outcomes, but the significantly higher prevalence of *pagC* in the seropathotypes associated with severe diseases indicates that this gene has some association with severe clinical outcome whether a strain carries the gene in OI-122 or in some other PAIs.

The distribution of PAI virulence genes and the phylogenetic analysis of *iha* and *pagC* support the hypothesis that OI-122, OI-43/48, and OI-57 are primarily associated with *eae*-positive strains in STEC. However, some *eae*-negative STEC serotypes, for example, O113:H21 and O91:H21, are also associated with life-threatening diseases such as HUS (5). Virulence factors such as subtilase cytotoxin AB5 (subAB5) and Saa (STEC autoagglutinating adhesion) are more commonly associated with *eae*-negative STEC. Moreover, it has been shown that some LEE-negative STEC strains, especially O113:H21, can invade tissue culture cells (38). Whole-genome comparison between nine *eae*-negative and five *eae*-positive STEC strains revealed that *eae*-negative strains did not carry any LEE-encoded effectors or other phage-encoded non-LEE effectors (39). These observations indicate that some differences in pathogenesis mechanisms may exist between *eae*-positive and *eae*-negative STEC strains. Additional studies, especially genomics and proteomics, are needed to determine the difference in the pathogenicity mechanisms between *eae*-negative and *eae*-positive STEC strains.

The strong association of OI-122, OI-57, and OI-43 with *eae*-positive STEC offers an important basis for STEC molecular risk

assessment (MRA). MRA, which uses 14 non-LEE-encoded virulence factors to distinguish high-risk from low-risk non-O157 STEC, was proposed by Coombes et al. in 2008 (13). Other researchers adopted this concept and applied it to their own studies (40–43). However, we demonstrated here that some of non-LEE-encoded effectors (*nleB*, *nle2-3*, *nleG5-2*, and *nleG6-2*) were primarily associated with *eae*-positive STEC strains. In addition, Mundy et al. (44) reported that *nleA* was present in 37 of 43 (86%) *eae*-positive STEC clinical strains but absent in 50 *eae*-negative STEC clinical strains. Konczy et al. (9) reported that *nleB* and *nleE* of OI-122 were highly correlated with LEE. Moreover, comparative genomics analysis demonstrated that all known phage-encoded non-LEE effector genes were absent in *eae*-negative STEC (39). Based on the MRA framework, which uses non-LEE effector genes as sole markers, all *eae*-negative virulence STEC strains, including HUS-associated O113:H21, O91:H21, and O104:H21, would be categorized as harmless STEC; other serotypes, for example, O103:H11 and O119:H25, which have not been reported to be associated with severe disease or outbreaks but carry non-LEE-encoded virulence effectors similar to those of O157 EHEC, would be considered outbreak- and severe disease-associated serotypes. Therefore, additional markers or methods of assessment, especially for *eae*-negative STEC, are needed to accurately distinguish highly pathogenic STEC from low-virulence or harmless STEC.

In summary, O-122 and OI-57 and their virulence genes were highly associated with seropathotypes that cause severe diseases and outbreaks. In addition, *ureC* and *terC*, located at OI-43/48, were also identified as markers related to high-risk seropathotypes. Virulence genes in PAIs that are associated with severe diseases can be used as markers to identify potentially highly virulent STEC. Furthermore, we demonstrated here that OI-122, OI-43/48, and OI-57 are highly associated with *eae*-positive STEC, which offers an important basis for STEC MRA.

ACKNOWLEDGMENTS

The study was supported in part by the Joint Institute for Food Safety and Applied Nutrition, University of Maryland, College Park, MD.

We thank Julie Kase, FDA/CFSAN, for providing DNA from positive controls for *stx* and *eae* subtyping.

REFERENCES

1. **Bettelheim KA.** 2007. The non-O157 Shiga-toxigenic (verocytotoxigenic) *Escherichia coli*: under-rated pathogens. Crit. Rev. Microbiol. **33**:67–87.
2. **Blanco M, Blanco JE, Mora A, Dahbi G, Alonso MP, Gonzalez EA, Bernardez MI, Blanco J.** 2004. Serotypes, virulence genes, and intimin types of Shiga toxin (verotoxin)-producing *Escherichia coli* isolates from cattle in Spain and identification of a new intimin variant gene (eae-xi). J. Clin. Microbiol. **42**:645–651.
3. **Scallan E, Hoekstra RM, Angulo FJ, Tauxe RV, Widdowson MA, Roy SL, Jones JL, Griffin PM.** 2011. Foodborne illness acquired in the United States: major pathogens. Emerg. Infect. Dis. **17**:7–15.
4. **Coombes BK, Gilmour MW, Goodman CD.** 2011. The evolution of virulence in non-O157 Shiga toxin-producing *Escherichia coli*. Front. Microbiol. **2**:90.
5. **Karmali MA, Mascarenhas M, Shen S, Ziebell K, Johnson S, Reid-Smith R, Isaac-Renton J, Clark C, Rahn K, Kaper JB.** 2003. Association of genomic O island 122 of *Escherichia coli* EDL 933 with verocytotoxin-producing *Escherichia coli* seropathotypes that are linked to epidemic and/or serious disease. J. Clin. Microbiol. **41**:4930–4940.
6. **Gal-Mor O, Finlay BB.** 2006. Pathogenicity islands: a molecular toolbox for bacterial virulence. Cell. Microbiol. **8**:1707–1719.
7. **Wickham ME, Lupp C, Mascarenhas M, Vazquez A, Coombes BK, Brown NF, Coburn BA, Deng W, Puente JL, Karmali MA, Finlay BB.** 2006. Bacterial genetic determinants of non-O157 STEC outbreaks and hemolytic-uremic syndrome after infection. J. Infect. Dis. **194**:819–827.
8. **Perna NT, Plunkett G, III, Burland V, Mau B, Glasner JD, Rose DJ, Mayhew GF, Evans PS, Gregor J, Kirkpatrick HA, Posfai G, Hackett J, Klink S, Boutin A, Shao Y, Miller L, Grotbeck EJ, Davis NW, Lim A, Dimalanta ET, Potamousis KD, Apodaca J, Anantharaman TS, Lin J, Yen G, Schwartz DC, Welch RA, Blattner FR.** 2001. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. Nature **409**:529–533.
9. **Konczy P, Ziebell K, Mascarenhas M, Choi A, Michaud C, Kropinski AM, Whittam TS, Wickham M, Finlay B, Karmali MA.** 2008. Genomic O island 122, locus for enterocyte effacement, and the evolution of virulent verocytotoxin-producing *Escherichia coli*. J. Bacteriol. **190**:5832–5840.
10. **Abu-Median AB, van Diemen PM, Dziva F, Vlisidou I, Wallis TS, Stevens MP.** 2006. Functional analysis of lymphostatin homologues in enterohaemorrhagic *Escherichia coli*. FEMS Microbiol. Lett. **258**:43–49.
11. **Taylor DE, Rooker M, Keelan M, Ng LK, Martin I, Perna NT, Burland NT, Blattner FR.** 2002. Genomic variability of O islands encoding tellurite resistance in enterohemorrhagic *Escherichia coli* O157:H7 isolates. J. Bacteriol. **184**:4690–4698.
12. **Yin X, Wheatcroft R, Chambers JR, Liu B, Zhu J, Gyles CL.** 2009. Contributions of O island 48 to adherence of enterohemorrhagic *Escherichia coli* O157:H7 to epithelial cells in vitro and in ligated pig ileal loops. Appl. Environ. Microbiol. **75**:5779–5786.
13. **Coombes BK, Wickham ME, Mascarenhas M, Gruenheid S, Finlay BB, Karmali MA.** 2008. Molecular analysis as an aid to assess the public health risk of non-O157 Shiga toxin-producing *Escherichia coli* strains. Appl. Environ. Microbiol. **74**:2153–2160.
14. **Imamovic L, Tozzoli R, Michelacci V, Minelli F, Marziano ML, Caprioli A, Morabito S.** 2010. OI-57, a genomic island of *Escherichia coli* O157, is present in other seropathotypes of Shiga toxin-producing *E. coli* associated with severe human disease. Infect. Immun. **78**:4697–4704.
15. **Wu B, Skarina T, Yee A, Jobin MC, Dileo R, Semesi A, Fares C, Lemak A, Coombes BK, Arrowsmith CH, Singer AU, Savchenko A.** 2010. NleG type 3 effectors from enterohaemorrhagic *Escherichia coli* are U-box E3 ubiquitin ligases. PLoS Pathog. **6**:e1000960. doi:10.1371/journal.ppat.1000960.
16. **Schubert S, Darlu P, Clermont O, Wieser A, Magistro G, Hoffmann C, Weinert K, Tenaillon O, Matic I, Denamur E.** 2009. Role of intraspecies recombination in the spread of pathogenicity islands within the *Escherichia coli* species. PLoS Pathog. **5**:e1000257. doi:10.1371/journal.ppat.1000257.
17. **Benedek O, Schubert S.** 2007. Mobility of the *Yersinia* high-pathogenicity island (HPI): transfer mechanisms of pathogenicity islands (PAIs) revisited (a review). Acta Microbiol. Immunol. Hung. **54**:89–105.
18. **Toma C, Martínez Espinosa E, Song T, Miliwebsky E, Chinen I, Iyoda S, Iwanaga M, Rivas M.** 2004. Distribution of putative adhesins in different seropathotypes of Shiga toxin-producing *Escherichia coli*. J. Clin. Microbiol. **42**:4937–4946.
19. **Beutin L, Miko A, Krause G, Pries K, Haby S, Steege K, Albrecht N.** 2007. Identification of human-pathogenic strains of Shiga toxin-producing *Escherichia coli* from food by a combination of serotyping and molecular typing of Shiga toxin genes. Appl. Environ. Microbiol. **73**:4769–4775.
20. **Tramuta C, Robino P, Oswald E, Nebbia P.** 2008. Identification of intimin alleles in pathogenic *Escherichia coli* by PCR-restriction fragment length polymorphism analysis. Vet. Res. Commun. **32**:1–5.
21. **Zheng J, Cui S, Teel LD, Zhao S, Singh R, O'Brien AD, Meng J.** 2008. Identification and characterization of Shiga toxin type 2 variants in *Escherichia coli* isolates from animals, food, and humans. Appl. Environ. Microbiol. **74**:5645–5652.
22. **Xia X, Meng J, McDermott PF, Ayers S, Blickenstaff K, Tran TT, Abbott J, Zheng J, Zhao S.** 2010. Presence and characterization of Shiga toxin-producing *Escherichia coli* and other potentially diarrheagenic *E. coli* strains in retail meats. Appl. Environ. Microbiol. **76**:1709–1717.
23. **Ju W, Shen J, Li Y, Toro MA, Zhao S, Ayers S, Najjar MB, Meng J.** 2012. Non-O157 Shiga toxin-producing *Escherichia coli* in retail ground beef and pork in the Washington D.C. area. Food Microbiol. **32**:371–377.
24. **Nakano M, Iida T, Ohnishi M, Kurokawa K, Takahashi A, Tsukamoto T, Yasunaga T, Hayashi T, Honda T.** 2001. Association of the urease gene with enterohemorrhagic *Escherichia coli* strains irrespective of their serogroups. J. Clin. Microbiol. **39**:4541–4543.
25. **Karch H, Schubert S, Zhang D, Zhang W, Schmidt H, Olschläger T,**

Hacker J. 1999. A genomic island, termed high-pathogenicity island, is present in certain non-O157 Shiga toxin-producing *Escherichia coli* clonal lineages. Infect. Immun. **67**:5994–6001.
26. **Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S.** 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum-parsimony methods. Mol. Biol. Evol. **28**:2731–2739.
27. **Nicholls L, Grant TH, Robins-Browne RM.** 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.
28. **Klapproth JM, Scaletsky IC, McNamara BP, Lai LC, Malstrom C, James SP, Donnenberg MS.** 2000. A large toxin from pathogenic *Escherichia coli* strains that inhibits lymphocyte activation. Infect. Immun. **68**:2148–2155.
29. **Klapproth JM, Sasaki M, Sherman M, Babbin B, Donnenberg MS, Fernandes PJ, Scaletsky IC, Kalman D, Nusrat A, Williams IR.** 2005. *Citrobacter rodentium lifA/efa1* is essential for colonic colonization and crypt cell hyperplasia in vivo. Infect. Immun. **73**:3196.
30. **Kelly M, Hart E, Mundy R, Marches O, Wiles S, Badea L, Luck S, Tauschek M, Frankel G, Robins-Browne RM, Hartland EL.** 2006. Essential role of the type III secretion system effector NleB in colonization of mice by *Citrobacter rodentium*. Infect. Immun. **74**:2328–2337.
31. **Steyert SR, Rasko DA, Kaper JB.** 2011. Functional and phylogenetic analysis of *ureD* in Shiga toxin-producing *Escherichia coli*. J. Bacteriol. **193**:875–886.
32. **Steyert SR, Kaper JB.** 2012. Contribution of urease to colonization by Shiga toxin-producing *Escherichia coli*. Infect. Immun. **80**:2589–2600.
33. **Friedrich AW, Lukas R, Mellmann A, Kock R, Zhang W, Mathys W, Bielaszewska M, Karch H.** 2006. Urease genes in non-O157 Shiga toxin-producing *Escherichia coli*: mostly silent but valuable markers for pathogenicity. Clin. Microbiol. Infect. **12**:483–486.
34. **Schmidt H, Zhang WL, Hemmrich U, Jelacic S, Brunder W, Tarr PI, Dobrindt U, Hacker J, Karch H.** 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.
35. **Newton HJ, Sloan J, Bulach DM, Seemann T, Allison CC, Tauschek M, Robins-Browne RM, Paton JC, Whittam TS, Paton AW, Hartland EL.** 2009. Shiga toxin-producing *Escherichia coli* strains negative for locus of enterocyte effacement. Emerg. Infect. Dis. **15**:372–380.
36. **Shen S, Mascarenhas M, Rahn K, Kaper JB, Karmali MA.** 2004. Evidence for a hybrid genomic island in verocytotoxin-producing *Escherichia coli* CL3 (serotype O113:H21) containing segments of EDL933 (serotype O157:H7) O islands 122 and 48. Infect. Immun. **72**:1496–1503.
37. **Girardeau JP, Bertin Y, Martin C.** 2009. Genomic analysis of the PAI ICL3 locus in pathogenic LEE-negative Shiga toxin-producing *Escherichia coli* and *Citrobacter rodentium*. Microbiology **155**:1016–1027.
38. **Luck SN, Badea L, Bennett-Wood V, Robins-Browne R, Hartland EL.** 2006. Contribution of FliC to epithelial cell invasion by enterohemorrhagic *Escherichia coli* O113:H21. Infect. Immun. **74**:6999–7004.
39. **Steyert SR, Sahl JW, Fraser CM, Teel LD, Scheutz F, Rasko DA.** 2012. Comparative genomics and *stx* phage characterization of LEE-negative Shiga toxin-producing *Escherichia coli*. Front. Cell. Infect. Microbiol. **2**:133.
40. **Bugarel M, Beutin L, Fach P.** 2010. Low-density macroarray targeting non-locus of enterocyte effacement effectors (*nle* genes) and major virulence factors of Shiga toxin-producing *Escherichia coli* (STEC): a new approach for molecular risk assessment of STEC isolates. Appl. Environ. Microbiol. **76**:203–211.
41. **Bugarel M, Beutin L, Martin A, Gill A, Fach P.** 2010. Micro-array for the identification of Shiga toxin-producing *Escherichia coli* (STEC) seropathotypes associated with hemorrhagic colitis and hemolytic uremic syndrome in humans. Int. J. Food Microbiol. **142**:318–329.
42. **Bugarel M, Martin A, Fach P, Beutin L.** 2011. Virulence gene profiling of enterohemorrhagic (EHEC) and enteropathogenic (EPEC) *Escherichia coli* strains: a basis for molecular risk assessment of typical and atypical EPEC strains. BMC Microbiol. **11**:142. doi:10.1186/1471-2180-11-142.
43. **Bosilevac JM, Koohmaraie M.** 2011. Prevalence and characterization of non-O157 Shiga toxin-producing *Escherichia coli* isolates from commercial ground beef in the United States. Appl. Environ. Microbiol. **77**:2103–2112.
44. **Mundy R, Jenkins C, Yu J, Smith H, Frankel G.** 2004. Distribution of *espI* among clinical enterohaemorrhagic and enteropathogenic *Escherichia coli* isolates. J. Med. Microbiol. **53**:1145–1149.