

Insertion site of the locus of enterocyte effacement in enteropathogenic and enterohemorrhagic *Escherichia coli* differs in relation to the clonal phylogeny of the strains

Lothar H. Wieler^{a,b,*}, Timothy K. McDaniel^{a,c}, Thomas S. Whittam^d,
James B. Kaper^a

^a Center for Vaccine Development, School of Medicine, University of Maryland, 685 W. Baltimore St., Baltimore, MD 21201, USA

^b Institut für Hygiene und Infektionskrankheiten der Tiere, Justus-Liebig-Universität Giessen, Frankfurter Str. 89, 35392 Giessen, Germany

^c Department of Microbiology and Immunology, Stanford University, School of Medicine, Stanford, CA 94305, USA

^d Institute of Molecular Evolutionary Genetics, The Pennsylvania State University, University Park, PA 16802, USA

Received 5 August 1997; accepted 29 August 1997

Abstract

The locus of enterocyte effacement pathogenicity island confers the attaching and effacing histopathology on epithelial cells infected with enteropathogenic and enterohemorrhagic *Escherichia coli*. We investigated the site of insertion of the locus of enterocyte effacement in *E. coli* strains in relation to their evolution based on conservation of housekeeping proteins in these strains. The results indicate that the insertion site of the locus of enterocyte effacement varies according to the evolutionary lineage, suggesting that it has inserted at multiple times and sites during the evolution of these pathogens.

Keywords: Pathogenicity island; Locus of enterocyte effacement; Enteropathogenic *Escherichia coli*; Enterohemorrhagic *Escherichia coli*; Evolution

1. Introduction

Recently, the 35-kb locus of enterocyte effacement (LEE) was reported to be crucially involved in the formation of the attaching and effacing lesion (AE lesion) on epithelial cells [1–3]. The LEE is present in all enteropathogenic *Escherichia coli* (EPEC) strains and in most enterohemorrhagic *E. coli* (EHEC) strains, including *E. coli* O157:H7. The considerably

lower G+C content (38%) of the LEE compared to the average *E. coli* genome (51%) indicates that the LEE was probably acquired by horizontal gene transfer from another species. In EPEC strain E2348/69 (O127:H6) and EHEC O157:H7 the LEE inserted into the selenocystyl tRNA locus (*selC*) [4]. *SelC* is considered a hot spot of insertion, since a pathogenicity island of uropathogenic *E. coli* also inserted into this locus [5].

The LEE is another example of so-called pathogenicity islands (PAI). PAIs are large, presumably horizontally acquired, multigene chromosomal DNA regions, encoding virulence properties which dis-

* Corresponding author.
Tel.: +49 (641) 99 38303; Fax: +49 (641) 99 38309;
E-mail: lothar.h.wieler@vetmed.uni-giessen.de

tinguish pathogenic from related nonpathogenic organisms [5,6]. PAIs have been reported in several other pathogens, including uropathogenic *E. coli*, *Salmonella*, *Helicobacter* and *Yersinia* [4-10]. Often such reports have examined only one strain of a group of pathogens. Whether the same insertion site is utilized for all members of a pathogenic species is not known, but a recent report showed that PAIs can be inserted at different chromosomal sites in different strains of uropathogenic *E. coli* [9].

In this study we have examined a larger number of diarrheagenic *E. coli* strains in order to determine the site of insertion of the LEE and to study whether the insertion of this island occurred at a single time in the evolution of this pathogen or at multiple times. To examine this issue, we used a well characterized set of *E. coli* strains for which evolutionary information was available. By multilocus enzyme electrophoresis of 20 housekeeping proteins of more than 1300 *E. coli* strains associated with enteric disease, 23 distinctive clones had previously been identified [11]. These clones are referred to as the DEC (diarrheagenic *E. coli*) [11]. A phylogenetic tree of the 15 most common DEC clones was constructed. Of these 15, the *E. coli* attaching and effacing gene (*eae*), indicative of the LEE, was found in nine clones. The EPEC and

EHEC clones harboring the LEE were furthermore clustered into four distinctive phylogenetic disease clusters. Disease cluster EPEC1 consists of classical serotypes such as O55:H6 and O127:H6. EPEC2 contains strains of serovars O111:H2 and O128:H2. The number of EHEC clustered so far is smaller. Strains of O111:H8 and O26:H11 form group EHEC2 and O157:H7 strains form EHEC1 [11,12].

2. Materials and methods

2.1. Bacterial strains

EPEC and EHEC strains were defined by serogroup, AE ability and by using DNA probes specific for the classes of pathogenic *E. coli* [2,11].

2.2. PCR

PCR reactions contained 5.0 µl template DNA (50 µl overnight broth+150 µl distilled water, 100°C, 10 min), 2.5 µl 10×PCR buffer, 1 µl (1 µM) of both primers [K255 (5'-GGTTGAGTCGATTGATCTCTGG-3'), K260 (5'-GAGCGAATATTCCGA-

Table 1

Results of PCR analysis of 34 LEE harboring DEC strains, distinguishing an intact *selC* locus from one disrupted by insertion of the LEE

Disease cluster	DEC clone	Serovar	No. of strains tested	PCR results with primer pairs				<i>selC</i> disrupted by LEE
				K295-K296	K255-K260	K295-K260	K261-K260	
EPEC1	E2348/69*	O127:H6	1	+	+	-	-	yes
EPEC1	1	O55:H6	4	+	+	-	-	yes
EPEC1	2	O55:H6	4	+	+	-	-	yes
EHEC1	3	O157:H7	4	+	+	-	-	yes
EHEC1	4	O157:H7	4	+	+	-	-	yes
(EPEC)	5	O55:H7	2	+	+	-	-	yes
EPEC2	11	O45:H2	1	-	-	+	+	no
		O128:H2	3	-	-	+	+	no
EPEC2	12	O111:NM	1	-	-	+	+	no
		O111:HN	1	-	-	+	+	no
		O111:H2	2	-	-	+	+	no
EHEC2	8	O111:NM	1	-	-	+	+	no
		O111:H8	2	-	-	+	+	no
		O111:H11	1	-	-	+	+	no
EHEC2	9	O26:HN	1	-	-	+	+	no
		O26:H11	3	-	-	+	+	no

*Strain E2348/69 is the prototype EPEC strain in which the insertion of the LEE into the *selC* locus was first described [7].

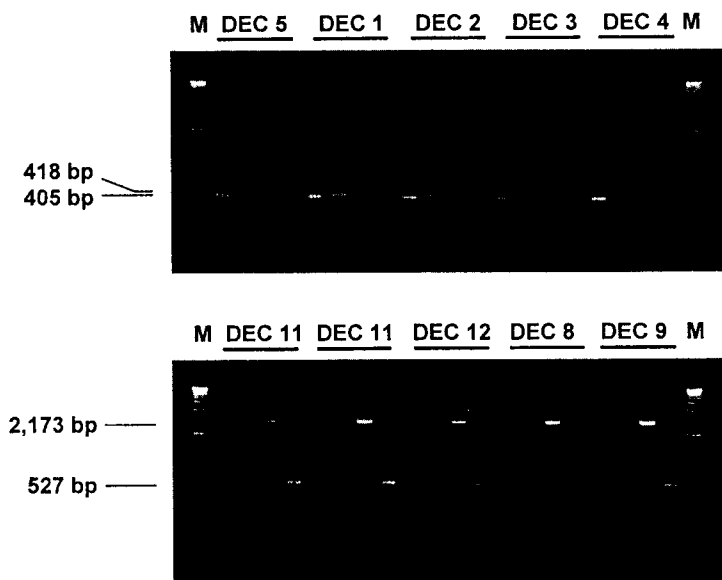


Fig. 1. PCR analysis of 10 representative *E. coli* strains from nine different DEC clones. Amplicons were size-fractionated by electrophoresis through 1.5% agarose gels. Amplicons with primer pairs K295-K296 (405 bp) and K255-K260 (418 bp) indicate a *selC* locus disrupted by the LEE, while amplicons with primer pairs K295-K260 (2173 bp) and K261-K260 (527 bp) indicate an intact *selC* locus. Products of all four PCRs are depicted in four consecutive lanes each. Lanes 1 and 22: molecular mass standard (M). Upper gel: lanes 2–5: strain O55:H7 (DEC 5); lanes 6–9: strain O55:H6 (DEC 1); lanes 10–13: strain O55:H6 (DEC 2); lanes 14–17: strain O157:H7 (DEC 3); lanes 18–21: strain O157:H7 (DEC 4). Lower gel: lanes 2–5: strain O45:H2 (DEC 11); lanes 6–9: strain O128:H2 (DEC 11); lanes 10–13: strain O111:H2 (DEC 12); lanes 14–17: strain O111:H8 (DEC 8); lanes 18–21: strain O26:H11 (DEC 9).

TATCTGGTT-3'), K261 (5'-CCTTGCAAATAAA-CACGGCGCAT-3'), K295 (5'-CGCCGATTTT-TCTTAGCCCA-3'), or K296 (5'-CATTCTGAAACAACTGCTC-3')], 200 μ M dNTPs and 1.5 units DNA polymerase (*Taq* DNA polymerase, BRL). Amplification was performed on a thermal minicycler MJ Research (Cambridge, MA): 5 min at 94°C, followed by 30 cycles of 1.30 min at 94°C, 1.30 min at 60°C, and 2 min at 72°C, followed by 5 min at 72°C.

2.3. Colony blot hybridization

Colony hybridization was done as in [2] with ³²P-labeled probes LEE A, B, C, and D, respectively.

2.4. DNA sequencing

DNA sequence analyses of amplicons K295-K296, K255-K260, K295-K260, and K261-K260, respectively, were performed using the prism ready reaction dye-deoxy terminator cycle sequencing kit (ABI, Ap-

plied Biosystems, Weiterstadt, Germany) according to the manufacturer's instructions. PCR products were purified with the Qiaquick-spin PCR purification kit (Qiagen, Hilden, Germany), and 100 ng purified DNA was used as template for each sequencing reaction.

3. Results and discussion

In the present work, we investigated the site of insertion of the LEE in different diarrheagenic *E. coli* by analyzing 34 *E. coli* strains from the nine LEE containing DEC clones by PCR. Colony blot hybridization with LEE probes A, B, C, D, spanning the entire 35-kb LEE region [2], showed that the sequences of the LEE are highly conserved in all 34 DEC strains (data not shown). Based on the data of McDaniel et al. [2] an intact *selC* can be distinguished from one disrupted by the LEE by the use of four different PCR primer pairs. Primer pair K295-K296 will amplify the left junction of the

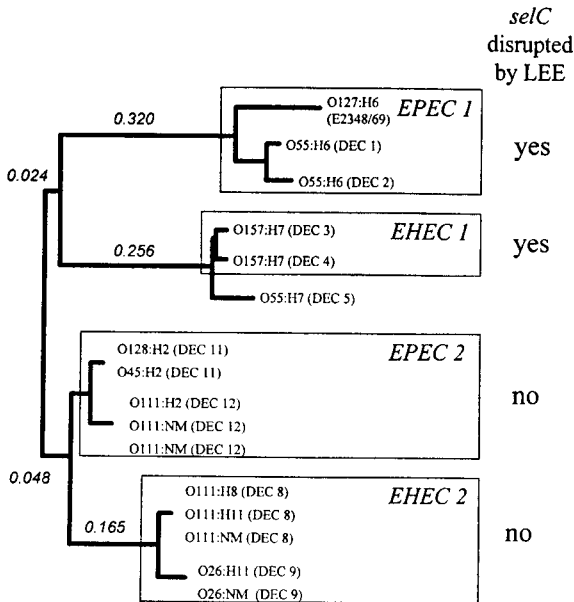


Fig. 2. Dendrogram based on multilocus enzyme electrophoresis of 20 housekeeping proteins. The dendrogram shows the clonal phylogeny of EPEC and EHEC groups of the nine LEE harboring DEC clones investigated.

LEE inserted into *selC* and yield a 405-bp amplicon while primer pair K255-K260 will amplify the right junction of the LEE inserted into *selC* and produce a 418-bp amplicon. In contrast, if the LEE insertion site is not within *selC* and the *selC* locus is intact, primer pairs K295-K260 and K261-K260 will give amplicons of 2173 and 527 bp, respectively (see Fig. 1). PCR results are listed in Table 1.

Strains of DEC groups 1–5, consisting of EPEC O55:H7 (two strains), EPEC1 serovar O55:H6 (eight strains), and EHEC1 serovar O157:H7 (eight strains), were shown to have a *selC* locus that was disrupted by the LEE. Strains of DEC groups 6–9, consisting of EPEC O55:H7 (two strains), EPEC1 serovar O55:H6 (eight strains), and EHEC1 serovar O157:H7 (eight strains), were shown to have a *selC* locus that was disrupted by the LEE. To determine the exact LEE insertion site in these strains, we analyzed the DNA sequence of the respective PCR fragments from one strain of each DEC clone. The DNA sequence analysis revealed that the LEE had inserted at the exact same site in each of these strains. The K-12 DNA sequences were 100% iden-

tical to those published by McDaniel et al. [2] (data not shown).

In contrast, strains of DEC groups 8, 9, 11, and 12 gave just the opposite PCR results. The PCR results indicate that in EPEC2 serovars O45:H2, O128:H2, O111:NM, O111:HN, and O111:H2, *selC* is intact. The same is true for EHEC2 serotypes O111:NM, O111:H8, O111:H11, O26:HN and O26:H11. These data indicate that in these strains, the LEE is inserted at a site other than *selC*. The location of the alternative site or sites is currently unknown. The alternative insertion sites are likely to be within tRNA genes as such genes are common, but not universal, insertion sites for pathogenicity islands in different pathogens [5,13]. Furthermore, in uropathogenic *E. coli* strains 536 [4] and J96 [9] and in *Salmonella* [10] PAIs were also shown to have inserted into various tRNA loci.

The clonal phylogeny of EPEC and EHEC strains based on multilocus enzyme electrophoresis results of housekeeping proteins is shown in Fig. 2. Recent data suggest that EHEC1 O157:H7 has emerged from an EPEC O55:H7-like progenitor that had already acquired the LEE [11]. Therefore, it is likely that the LEE moved into *selC* in a common EHEC1 and EPEC1 progenitor. As the exact LEE insertion site in strains of EHEC2 and EPEC2 is not yet defined, two possible scenarios in their evolution can be envisioned. These strains may have one common progenitor, which must have emerged differently from the EHEC1 and EPEC1 progenitor. This progenitor must have emerged more recently, because EHEC2 and EPEC2 show a lower level of clonal diversity, which would require less time to accumulate. Alternatively, both disease clusters may have developed independently. Strains of EPEC2 mostly harbor flagellar antigen 2 (H type 2). In EPEC strains, H types seem to be more conserved among strains of the same cluster than somatic antigens (O types). The H types of EHEC2 strains (H types 8 and 11) differ from H type 2. Therefore it is more likely that EHEC2 and EPEC2 developed independently.

These studies provide insight into the evolution of EPEC and EHEC strains. We are currently investigating the insertion site of the LEE in EPEC2 and EHEC2 strains to see if the LEE inserted at a single site, presumably during a single event, or if insertion

of this pathogenicity island occurred at numerous times into various sites during the evolution of these important enteric pathogens.

Acknowledgments

Supported by NIH Grant AI 21657.

References

- [1] Jarvis, K.G., Giron, J., Jerse, A.E., McDaniel, T.K., Donnenberg, M. and Kaper, J.B. (1995) Enteropathogenic *Escherichia coli* contains a putative type III secretion system necessary for the export of proteins involved in attaching and effacing lesion formation. Proc. Natl. Acad. Sci. USA 92, 7996–8000.
- [2] McDaniel, T.K., Jarvis, K.J., Donnenberg, M.S. and Kaper, J.B. (1995) A genetic locus of enterocyte effacement conserved among diverse enterobacterial pathogens. Proc. Natl. Acad. Sci. USA 92, 1664–1668.
- [3] McDaniel, T.K. and Kaper, J.B. (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.
- [4] Blum, G., Ott, M., Lischewski, A., Ritter, A., Imrich, H., Tschäpe, H. and Hacker, J. (1994) Excision of large DNA regions termed pathogenicity islands from tRNA-specific loci in the chromosome of an *Escherichia coli* wild-type pathogen. Infect. Immun. 62, 606–614.
- [5] Lee, C.A. (1996) Pathogenicity islands and the evolution of bacterial pathogens. Infect. Agents Dis. 5, 1–7.
- [6] Mecsas, J. and Strauss, E.J. (1996) Molecular mechanisms of bacterial virulence: type III secretion and pathogenicity islands. Emerg. Infect. Dis. 2, 271–288.
- [7] Censini, S., Lange, Ch., Xiang, Z., Crabtree, J.E., Ghiara, P., Borodovsky, M., Rappuoli, R. and Covacci, A. (1996) *cag*, a pathogenicity island of *Helicobacter pylori*, encodes type I-specific and disease-associated virulence factors. Proc. Natl. Acad. Sci. USA 93, 14648–14653.
- [8] Hacker, J., Bender, L., Ott, M., Wingender, J., Lund, B., Marre, R. and Goebel, W. (1990) Deletions of chromosomal regions coding for fimbriae and hemolysins occur in vitro and in vivo in various extraintestinal *Escherichia coli* isolates. Microb. Pathogen. 8, 213–225.
- [9] Swenson, D.L., Bukanov, N.O., Berg, D.E. and Welch, R.A. (1996) Two pathogenicity islands in uropathogenic *Escherichia coli* J96: cosmid cloning and sample sequencing. Infect. Immun. 64, 3736–3743.
- [10] Hensel, M., Shea, J.E., Bäumlner, A.J., Gleeson, C., Blattner, F. and Holden, D.W. (1997) Analysis of the boundaries of *Salmonella* pathogenicity island 2 and the corresponding chromosomal region of *Escherichia coli* K-12. J. Bacteriol. 179, 1105–1111.
- [11] Whittam, T.S., Wolfe, M.L., Wachsmuth, I.K., Ørskov, F., Ørskov, I. and Wilson, R.A. (1993) Clonal relationships among *Escherichia coli* strains that cause hemorrhagic colitis and infantile diarrhea. Infect. Immun. 61, 1619–1629.
- [12] Whittam, T.S. and McGraw, E.A. (1996) Clonal analysis of EPEC serogroups. Rev. Microbiol. 27 (Suppl. 1), 7–16.
- [13] Reiter, W.-D., Palm, P. and Yeats, S. (1989) Transfer RNA genes frequently serve as integration sites for procaryotic genetic elements. Nucleic Acids Res. 17, 1907–1914.