RESEARCH LETTER

Comparison between a bovine and a human enterohaemorrhagic *Escherichia coli* strain of serogroup O26 by suppressive subtractive hybridization reveals the presence of atypical factors in EHEC and EPEC strains

Marjorie Bardiau, Bernard Taminiau, Jean-Noël Duprez, Sabrina Labrozzo & Jacques G. Mainil

Department of Infectious and Parasitic Diseases, Bacteriology, Veterinary Faculty, University of Liège, Liège, Belgium

Correspondence: Marjorie Bardiau, Department of Infectious and Parasitic Diseases, Bacteriology, Veterinary Faculty, University of Liège, Liège, Belgium. Tel.: +32 4 366 40 52; fax: +32 4 366 42 63; e-mail: mbardiau@ulg.ac.be

Received 12 November 2011; revised 28 February 2012; accepted 5 March 2012. Final version published online 2 April 2012.

DOI: 10.1111/j.1574-6968.2012.02542.x

Editor: Ian Henderson

Keywords
enterohaemorrhagic *Escherichia coli*; suppressive subtractive hybridization; serogroup O26; PAI I_{CL3}.

Abstract

Enterohaemorrhagic *Escherichia coli* (EHEC) strains are responsible for food poisoning in humans in developed countries via consumption of vegetal and animal foodstuffs contaminated by ruminant feces. The clinical conditions caused by EHEC strains vary from undifferentiated diarrhea to hemorrhagic colitis with, in a few cases, the appearance of the hemolytic uremic syndrome, which can lead to death. Most EHEC strains can be found in the gut of healthy ruminants, but some of the strains, belonging to O26, O111, O118 serogroups, for example, are also responsible for digestive disorders in calves. The aim of this research was to study the genomic differences between two EHEC strains of serogroup O26 isolated from a young calf and a human with diarrhea, to identify specific sequences of the bovine strain that could be implicated in initial adherence or host specificity. No sequence implicated in host specificity was found during our study. Finally, several factors, not usually present in EHEC strains of serogroup O26, were identified in the bovine strain. One of them, the PAI I_{CL3} locus initially presented as a marker for LEE-negative VTEC strains, was found in 11.3% of EPEC and EHEC strains.

Introduction

In humans, enterohaemorrhagic *Escherichia coli* (EHEC) is responsible for the production of diarrhea, generally accompanied by hemorrhagic colitis with, in a few percent of cases, the occurrence of renal sequelae (hemolytic uremic syndrome), which can lead to death. EHEC strains were recognized as a distinct class of pathogenic *E. coli* in 1983 after two outbreaks in the United States (Wells *et al.*, 1983). Today, they represent a significant problem for public health in developed countries. Indeed, large outbreaks are caused by EHEC strains (Nataro & Kaper, 1998), and transmission often occurs via consumption of vegetal and animal foodstuffs contaminated by feces of adult ruminants (mainly cattle), which can be healthy carriers (Caprioli *et al.*, 2005). The most common EHEC serotype is O157:H7, which causes disease worldwide, but other serogroups such as O26, O111, and/or O103 are also of high epidemiological importance in some countries (Brooks *et al.*, 2005; Bettelheim, 2007). In the veterinary field, different serogroups of EHEC strains (O5, O26, O111, O118, for example) are directly associated with diarrhea in 2-week to 2-month-old calves (Moxley & Francis, 1986; Stordeur *et al.*, 2000; Hornitzky *et al.*, 2005). The consequences are economic losses owing to a delay in growth and weakness of the calves.

Some pathogenic *E. coli* are host specific, based upon the production of host-specific properties, in particular adhesins and colonization factors (for example, human typical EPEC, rabbit-EPEC, and porcine-VTEC). However, the actual situation about the host specificity regarding those EHEC serogroups (e.g. O26 and O111) infecting both humans and young calves, and present in healthy adult ruminants, is unknown: Do some isolates possess some degree of host specificity or can all isolates in fact infect all the hosts? As with host-specific

© 2012 Federation of European Microbiological Societies
Published by Blackwell Publishing Ltd. All rights reserved

FEMS MICROBIOLOGY LETTERS

pathogenic *E. coli*, the basis of any host specificity of those EHEC strains may be related to the production of specific colonization factors, although such adhesins of EHEC strains have not yet been identified (Bardiau *et al.*, 2009).

The aim of this study was (1) to explore the genomic differences, using suppressive subtractive hybridization (SSH), between two EHEC strains of serogroup O26, one isolated from a young calf and the other isolated from a human with diarrhea, to identify specific sequences of the bovine strain; (2) to analyze the bovine strain-specific sequences regarding their potential implication in adherence to epithelial cells; and (3) to study the prevalence of these strain-specific sequences in a collection of human and bovine EHEC and EPEC strains.

Materials and methods

Bacterial strains

Subtractive suppressive hybridization (SSH) was performed between the bovine EHEC strain 4276 of serogroup O26 isolated in Ireland from a diarrheic calf (Kerr *et al.*, 1999) and the human EHEC strain 11368 of serogroup O26 isolated in Japan from a human suffering from diarrhea (Ogura *et al.*, 2009). The distribution of the specific sequences was investigated in additional EHEC ($n = 44$) and EPEC ($n = 27$) strains of serogroup O26 isolated from humans ($n = 27$) and from cattle ($n = 44$). Most of the strains have been described previously (Szalo *et al.*, 2004; Bardiau *et al.*, 2009), and their characteristics are described in the supplemental Table S1.

Pulsed field gel electrophoresis (PFGE)

PFGE was performed as already described (Cobbaut *et al.*, 2009; Ooka *et al.*, 2009) on most of the tested strains. In brief, bacterial cells were embedded in 1.8% Certified Low Melt Agarose (Bio-Rad Laboratories, Inc., Tokyo, Japan), lysed with a buffer containing 0.2% sodium deoxycholate, 0.5% N-lauroylsarcosine, and 0.5% Brij-58, and treated with 100 μg mL^{-1} proteinase K. XbaI-digested genomic DNA was separated using CHEF MAPPER (Bio-Rad Laboratories, Inc.) with 1% Pulsed Field Certified Agarose (Bio-Rad Laboratories, Inc.) at 6.0 V cm^{-1} for 22 h and 18 min with pulsed times ranging from 47 to 44.69 s. Size of each DNA band was estimated by Biogene (Vilber Lourmat, France). The banding patterns were analyzed using the Dice coefficient, with an optimization and position tolerance of 1%. Dendrograms were prepared by the unweighted-pair group method using arithmetic average algorithm (UPGMA).

SSH

Genomic DNA was extracted from *E. coli* strain 4276 and *E. coli* strain 11368 using the cetyltrimethylammonium bromide procedure described by Ausubel *et al.* (1994). Subtractive hybridization was carried out using the PCR-Select Bacterial Genome Subtractive kit (Clontech) as recommended by the manufacturer. The bovine EHEC strain 4276 was the tester, and the human EHEC strain 11368 was the driver. The PCR products obtained were cloned into the pGEM-T Easy Vector System (Promega) and transformed into *E. coli* JM109. The recombinant clones were plated onto LB plates containing ampicillin (100 μg mL^{-1}), 0.2 mM IPTG (isopropyl-β-D-thiogalactopyranoside), and 40 μg X-Gal mL^{-1} (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside). White colonies containing recombinant plasmids with inserts were picked up, grown overnight at 37 °C in LB broth with ampicillin (100 μg mL^{-1}), and stored in a freezer (−20 °C) until further use.

DNA sequencing

The plasmid inserts were amplified by PCR with specific primers (nested primers 1 and 2R from the Clontech protocol), and the DNA fragments were purified using the NucleoSpin Extract II kit (Macherey-Nagel) according to the manufacturer's instructions. Sequencing of the two DNA strands was performed by the dideoxynucleotide triphosphate chain termination method with a 3730 ABI capillary sequencer and a BigDye Terminator kit version 3.1 (Applied Biosystems) at the GIGA (Groupe Interdisciplinaire de Génoprotéomique Appliquée, University of Liège, Belgium). Sequence analysis was performed using Vector NTI 10.1.1 (Invitrogen). DNA sequences were further examined for homology with the National Center for Biotechnology Information (NCBI) BLASTN and BLASTX programs (http://www.ncbi.nlm.nih.gov/BLAST/). The expectation value of 0.001 was chosen as the cutoff.

DNA colony hybridization

Several EHEC strain 4276–specific sequences were chosen for extended analysis. Their distribution was searched for in the collection of 71 bovine and human EHEC and EPEC strains by DNA colony hybridization at 65 °C on Whatman 541 paper filters (Whatman), as previously described (Mainil *et al.*, 1997). The DNA probes were derived by PCR from plasmid inserts obtained with SSH. The DNA probe fragments were purified using the NucleoSpin Extract II kit (Macherey-Nagel), according to the manufacturer's instructions, and labeled with α^{32}P-dCTP

© 2012 Federation of European Microbiological Societies
Published by Blackwell Publishing Ltd. All rights reserved

(Perkin-Elmer) by random priming using the Ready-To-Go dCTP-labeling beads (Amersham Biosciences). Labeled DNA probes were purified with Microcon-YM30 spin columns (Millipore).

PCR reactions

All primers and PCR conditions used in this study have been described previously (China *et al.*, 1996; Shen *et al.*, 2004; Durso *et al.*, 2005) (Supporting Information, Table S2). DNA extraction was carried out by a boiling method as described previously by China *et al.* (1996). For the PCR, the following mixture was used: 1 U of *Taq* DNA polymerase (New England Biolabs), 5 μL of 2 mM deoxynucleoside triphosphates, 5 μL of 10× ThermoPol Reaction Buffer (20 mM Tris–HCl (pH 8.8, 25 °C), 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% Triton X-100), 5 μL of each primer (10 μM), and 3 μL of a DNA template in a total volume of 50 μL.

Statistical analysis

A Fisher's exact test was performed to assess statistical differences ($P < 0.01$).

Results

PFGE

PFGE profiles were obtained for 60 of the 73 tested strains. Others strains did not present any restriction profile for XbaI or could not be tested. The 60 distinct electrophoresis profiles were used for dendrogram construction (Fig. S1). The dendrogram showed five clusters, assuming a cutoff of 45% of similarity. When a cutoff of over 80% of similarity was adopted, 38 different clusters were found, indicating the high genetic variability among the strains.

Identification of the bovine EHEC strain 4276–specific genes in the subtractive library

A total of 1920 clones resulting from the SSH process were obtained, of which 772 were randomly sequenced, resulting in 296 contigs after removal of redundant sequences. The specificity of the contigs to the bovine EHEC strain (strain 4276) was determined by a BLASTN search with the human EHEC strain (strain 11368) genome sequenced by Ogura *et al.* (2009). Of the 296 nonredundant DNA contigs, 115 contained genes different from those of the human EHEC strain (strain 11368).

Analysis of the bovine EHEC strain 4276–specific genes

BLASTN and BLASTX against the GenBank were searched for the 115 contigs specific to the bovine strain (Table 1 and Table S3). Several groups of genes were revealed by more than one clone: colicin resistance genes, multiple antibiotic resistance region from *Salmonella enterica*, phages P1 and P7, pathogenicity island (termed PAI I_{CL3}) described in the VTEC O113:H21 *E. coli* CL3 (containing putative adhesins and hemolysins), genes from the genomic islands GEI 3.21 described in *E. coli* O111:H−, transposase from *Enterobacter cloacae*, *E. coli* and *Acinetobacter baumanii*, predicted type I restriction-modification enzyme from *E. coli* 0127:H6 E2348/69, DEAD/DEAH box helicase from *Nitromonas europea*, SNF2 family helicase from *E. coli* strain E24377A, plasmid pO111_2 from *E. coli* O111:H−, and plasmid pSMS35_8 from *E. coli* SMS-3-5. BLASTN revealed six sequences that are not homologous to any annotated DNA sequences in GenBank. The other sequences were detected in only one clone and corresponded to genes specific to *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, *Citrobacter rotendium*, *Shigella sonnei*, *Erwinia* sp., *Desulfurispirillum indicum*, *Dickeya zeae*, *Pantoea ananatis*, and several strains of *E. coli.*

Distribution of specific sequences in a collection of EHEC and EPEC isolates

Several sequences (in bold in Table 1 and Table S3) were chosen for further characterization based upon the frequency of the contigs in the subtractive library or upon the putative involvement in adherence to the eukaryotic cells or in host specificity: genes from PAI I_{CL3}, four sequences with no homology, genes from P1 and P7 phages, genes from genomic island GEI 3.21, hypothetical proteins from E23477A strain, DEAD/DEAH box helicase from *Nitromonas* sp., genes from *E. coli O111:H−* strain 11128, transposase from *A. baumanii.*, ABC transporter from *D. zeae*, and *avrA* genes from *E. coli* strain CB769. The regions of DNA homologous to that previously identified in the subtractive library were searched for in EHEC and EPEC strains of serogroup O26 isolated from human and from cattle using DNA colony hybridization (Table 2) or using specific PCR for PAI I_{CL3} locus (Table 3).

Statistical analyses were performed to assess differences in the presence of the fragments according to host specificity (human or bovine) and/or pathotype (EHEC or EPEC). Two sequences, both homologous to the genomic island GEI 3.21 from *E. coli* O111:H−, were statistically

© 2012 Federation of European Microbiological Societies
Published by Blackwell Publishing Ltd. All rights reserved

Table 1. Results of the BLASTN against the GenBank searched for the 115 contigs specific to the bovine strains. Sequences in bold were chosen for further characterization

Function	Number of nonredundant and specific contigs	BLASTN results	Species
Antibiotic resistance	9	Colicin resistance	*Escherichia coli*
	3	Multiple antibiotic resistance region	*Salmonella enterica*, *Klebsiella pneumoniae*
Mobile functions	7	**Transposase**	*Enterobacter cloacae*, *E.*, *Acinetobacter baumanii*
	1	Excisionase	*E. coli*
Genomic island	1	Genomic island GEI1.94	*E. coli*
	1	Genomic island AGI-5	*E. coli*
	7	**Genomic island GEI3.21**	*E. coli*
Unknown function	14	**Hypothetical protein**	*E. coli*, *Pseudomonas aeruginosa*, *Citrobacter rodentium*
	6	**No homology**	–
Adherence related	7	**Pathogenicity island I (PAI I_{CL3})**	*E. coli*
	1	**Putative hemolysin/hemagglutinin**	*C. rodentium*
	1	espP	*E. coli*
	1	tonB	*Shigella sonnei*
Metabolism	3	**Predicted type I restriction-modification enzyme, S subunit**	*E. coli*
	2	N-6 DNA methylase	*Desulfurispirillum indicum*, *Dickeya zeae*
	1	Galactosyl transferase	*Erwinia* sp.
	4	**DEAD/DEAH box helicase**	*Nitrosomonas europaea*, *P. aeruginosa*
	3	SNF2 family helicase	*E. coli*
	2	**ABC transporter**	*Dickeya zeae*, *Nitrosomonas europaea*
Phage related	8	**Enterobacteria phage P7 or P1**	*Enterobacteria* phage P7 or P1
	1	Putative tail fiber assembly protein	*S. sonnei*
	1	**Putative phage repressor protein**	*E. coli*
Other	1	FhaB	*Pantoea ananatis*
	1	**avrA**	*E. coli*
	23	**Plasmid pO111_2**	*E. coli*
	1	Plasmid pCRP3	*C. rodentium*
	3	Plasmid pSMS35_8	*E. coli*
	1	Plasmid pHUSEC41-1	*E. coli*
	1	Plasmid pO145-NM	*E. coli*
Total	115		

associated with EPEC strains in comparison with EHEC strains. One of the fragments homologous to P1 phage was statistically associated with EHEC in comparison with EPEC. The sequence homologous to the predicted type I restriction-modification enzyme from *E. coli O127:H6* strain E2348/69 was statistically associated with strains isolated from humans in comparison with strains isolated from bovines. All the other fragments were associated with neither pathotype nor host.

Distribution of the PAI I_{CL3} locus in human and bovine EPEC and EHEC strains

Shen *et al.* (2004) first described the PAI I_{CL3} locus in the O113:H21 VTEC strain CL3. PAI I_{CL3} is a hybrid genomic region composed of genes similar to EDL933 (serotype O157:H7) O islands 122 and 48, *Yersinia pestis*, *Ralstonia solanacearum*, *Pseudomonas syringae*, *Fusobacterium nucleatum*, *Bacillus subtilis*, *S. enterica*, and *Sulfolobus tokodaii* (Table 3). To date, PAI I_{CL3} has been detected only in *eae*-negative VTEC strains associated with diseases in humans and never in any other pathogenic or commensal *E. coli*, and it may therefore be used as a new marker for those strains (Girardeau *et al.*, 2009). As several genes of PAI I_{CL3} have been identified here in the bovine EHEC strain 4276 of serogroup O26, their distribution was studied with specific PCRs in the collection of human and bovine EHEC and EPEC strains.

Eight strains (three human EPEC and five human and bovine EHEC strains) were found to be positive for several PCRs targeting different genes of the PAI I_{CL3} locus (Table 3). According to their PFGE pattern, these eight strains are not closely related. Indeed, they are present in the five clusters revealed by the PFGE dendrogram with a

© 2012 Federation of European Microbiological Societies
Published by Blackwell Publishing Ltd. All rights reserved

Table 2. Distribution of specific sequences resulting from SSH analyses

	Number of strains																			
		Number and percentage of strains found positive for																		
		LG100	Gamma D11	LG53	LG31	LG106	LG88	LG69	LG109	LG18	Epsilon G6	Alpha F3	Dzeta H6	*LG28*	*LG45*	*LG12*	Nu G3	LG2	Espilon E1	LG92	LG16
	No. of tested strains	No homo-logy	No homo-logy	No homo-logy	No homo-logy	Entero-bacteria phage P7	Recom-bination enhance-ment function	Plasmid pO111 _2	Plasmid pO111 _2	Putative phage repressor protein	Hypo-thetical protein	Hypo-thetical protein	Hypo-thetical protein	ABC trans-porter	DEAD/ DEAH box helicase	DEAD/ DEAH box heli-case	avrA	Geno-mic island GEI 3.21	Geno-mic island GEI 3.21	Predicted type I restriction-modification enzyme, S subunit	Trans-posase
Total no. of strains	75	3 (4%)	6 (8%)	52 (69%)	2 (3%)	14 (19%)	15 (20%)	14 (19%)	5 (7%)	12 (16%)	0 (0%)	2 (3%)	23 (31%)	1 (1%)	1 (1%)	4 (5%)	36 (48%)	24 (32%)	16 (21%)	12 (16%)	23 (31%)
Bovine strains	44	3 (7%)	4 (9%)	31 (70%)	1 (2%)	9 (20%)	10 (23%)	10 (23%)	1 (2%)	7 (16%)	0 (0%)	1 (2%)	11 (25%)	1 (2%)	1 (2%)	4 (9%)	25 (57%)	15 (34%)	11 (25%)	**3 (7%)**	13 30%)
Human strains	27	0 (0%)	2 (7%)	21 (78%)	1 (4%)	5 (19%)	5 (19%)	4 (15%)	4 (15%)	5 (19%)	0 (0%)	0 (0%)	12 (44%)	0 (0%)	0 (0%)	0 (0%)	11 (41%)	9 (33%)	5 (19%()	**9 (33%)**	10 (37%)
EPEC strains	27	0 (0%)	2 (7%)	19 (70%)	1 (4%)	2 (7%)	**1 (4%)**	2 (7%)	1 (4%)	2 (7%)	0 (0%)	1 (4%)	6 (22%)	0 (0%)	0 (0%)	0 (0%)	8 (30%)	**15 (56%)**	**12 (44%)**	4 (15%)	7 (26%)
EHEC strains	44	3 (7%)	4 (9%)	33 (75%)	1 (2%)	12 (27%)	**14 (32%)**	12 (27%)	4 (9%)	10 (23%)	0 (0%)	1 (2%)	17 (39%)	1 (2%)	1 (2%)	4 (9%)	28 (64%)	**9 (20%)**	**4 (9%)**	8 (18%)	16 (36%)

similarity of 45%, suggesting that these genes were horizontally acquired. No statistical difference was associated with the pathotype and/or the host origin ($P < 0.01$). This genomic island can in fact be divided into four parts: two genomic segments (GS-I inserted and GS-II including two genes of OI-122) bordered by OI-48 segments either side (Shen *et al.*, 2004). The eight strains were tested positive here with the PCRs for the three genes of GS-I and for all six genes of the two OI-48 segments. To verify whether Z1640 gene is intact or not, we performed two PCRs: one PCR targeting the Z1640-1 and Z1640-3 sequences (using Z1640-F and Z1640-R primers) and one PCR targeting the Z1640-1 and S1 sequences (using Z1640-F and S1-bis-R primers). The eight strains were positive only with the Z1640/S1 PCR. On the other hand, only the S4 gene of GS-II was detected in all eight strains, while the other genes (including *S10* and *S11* genes of OI-122) were detected in none to six strains only.

Discussion

Several serogroups of EHEC strains (e.g. O5, O26, O111, O118) can infect both humans and calves and can also be found in healthy cattle. Factors implicated in host specificity have been identified for some other pathogenic *E. coli* strains, but not for EHEC strains. Such factors could be based on proteins intervening in the colonization stage (adhesins, for example). Therefore, we wanted to explore the genomic differences between a bovine EHEC strain of serogroup O26 and a human EHEC strain of serogroup O26 using SSH to identify specific sequences of the bovine strain. This study aimed to explore the potential implication in initial adherence or host specificity of the specific sequences.

In the SSH library, we obtained 115 unique fragments that were specific to the bovine strain. These fragments include sequences with homology to genes or pathogenicity islands (PAIs) present only in other specific *E. coli* pathotypes (e.g. VTEC) or other species (e.g. *Klebsiella*, *Nitromonas*), which are not known to be present in EHEC strains of serogroup O26. This heterogeneity supports the hypothesis of a horizontal acquisition of genomic regions from other pathogenic bacteria (Brzuszkiewicz *et al.*, 2009; Juhas *et al.*, 2009; Kelly *et al.*, 2009). Moreover, it reflects the genomic plasticity of EHEC and/or *E. coli* strains. This finding supports the hypothesis of Mokady *et al.* (2005), suggesting that this variation in the genome contents of *E. coli* could indicate that its evolutionary strategy tends to create a mixed assortment of virulence factors coming from various pathogenic strains. This combination leads to a unique set of such factors, which helps the bacteria to better survive.

© 2012 Federation of European Microbiological Societies
Published by Blackwell Publishing Ltd. All rights reserved

Table 3. Distribution of the genes carried by the PAI I_{CL3} locus in the eight positive strains

			Strain	4276	TC6165	TC6166	02/057	99/147	A39	63	A14
			Pathotype	EHEC	EPEC	EPEC	EPEC	EHEC	EHEC	EHEC	EHEC
			Host	B	H	H	H	H	B	B	B
Similar protein (% identity)			Genes								
Z1635, unknown protein, *E. coli* EDL933 (97)	OI-48		Z1635	+	+	+	+	+	+	+	+
Z1636, unknown protein, *E. coli* EDL933 (96)			Z1636	+	+	+	+	+	+	+	+
Z1637, unknown protein, *E. coli* EDL933 (95)			Z1637	+	+	+	+	+	+	+	+
YPO2491, putative hemolysin activator, *Y. pestis* CO92 (64)	GSI		S1	+	+	+	+	+	+	+	+
RS02573, putative hemolysin			S2	+	+	+	+	+	+	+	+
activating-like protein, *R. solanacearum* (57)											
YPO2490, putative hemolysin (53)			S3	+	+	+	+	+	+	+	+
YPO0599, putative adhesin, *Y. pestis* CO92 (50)											
YPO2490, putative hemolysin (41)	GSII		S4	+	+	+	+	+	+	+	+
YPO0599, putative adhesin, *Y. pestis* CO92 (39)											
YPO0599, putative adhesin, *Y. pestis* CO92 (83)			S5	+	+	–	–	+	+	–	–
Y2435, putative transposase, *Y. pestis* KIM (38)			S6	+	+	–	–	+	+	+	+
TnpA, transposase, *P. syringae* (60)			S7	–	–	–	–	–	–	–	–
FN0835, hypothetical protein, *F. nucleatum* ATCC 25586 (27)			S8	–	–	–	–	–	–	–	–
YozI, unknown protein, *B. subtilis* (32)			S9	–	+	+	–	–	–	–	–
Z4322, unknown protein, *E. coli* EDL933 (94)		OI-122	S10	–	–	–	–	–	–	–	–
Z4321, unknown protein, *E. coli* EDL933 (98)			S11	–	+	+	–	–	–	+	+
Orf1, similarity with helicase, *S. enterica* (40)			S12	+	–	–	–	–	–	–	–
ST0071, hypothetical esterase *Sulfolobus tokodaii* (30)			S13	–	+	–	–	–	–	–	–
Y2679, hypothetical protein, *Y. pestis* KIM (39)			S14	–	–	–	–	–	+	+	+
Z1640, unknown protein, *E. coli* EDL933 (90)											
Z1641, unknown protein, *E. coli* EDL933 (96)	OI-48		S15	+	+	+	+	+	+	+	+
Z1642, unknown protein, *E. coli* EDL933 (99)											
Z1643, unknown protein, *E. coli* EDL933 (97)			Z1643	+	+	+	+	+	+	+	+
Z1644, unknown protein, *E. coli* EDL933 (98)			Z1644	+	+	+	+	+	+	+	+

B, bovine; H, human.

© 2012 Federation of European Microbiological Societies
Published by Blackwell Publishing Ltd. All rights reserved

The PAI I_{CL3} locus, first described by Shen *et al.* (2004) in the VTEC O113:H21 *E. coli* CL3, was found in 11.3% of the tested EHEC and EPEC strains of serogroup O26. These results are surprising when compared to those obtained by Girardeau *et al.* (2009), suggesting that PAI I_{CL3} is unique to LEE-negative VTEC strains and that this locus thus provides a new marker for such strains. We have reported here that the locus could also be present in *eae*-positive strains belonging to a major serogroup involved in human diseases. Girardeau *et al.* (2009) have suggested that PAI I_{CL3} used to be present in most *E. coli* pathotypes but that many of these pathotypes have undergone extensive deletions [probably via homologous recombination between insertion sequences (IS) elements, which removed almost the entire locus]. We can assume that our positive strains were not deleted for this locus. Another possible explanation is that these strains have recently acquired the PAI I_{CL3} locus via horizontal transfer, which hypothesis is supported by the fact that the PAI I_{CL3}-positive strains are not closely related.

Concerning host specificity, only one sequence appears to be statistically specific to human strains in comparison with bovine strains. Nevertheless, this sequence is only present in a few strains (7% of bovine strains and 33% of human strains) and therefore could not represent a host-specific marker. Moreover, three sequences were statistically associated with the pathotype (EHEC or EPEC), but these sequences were not present in more than half of the EPEC strains. However, host-specific factors could, perhaps, not be detected by SSH for one of the following reasons: (1) the subtraction is nonexhaustive and host-specific factors were not detected; (2) this host specificity is not based on the presence/absence of specific factors/genes; and (3) there is no host specificity.

In conclusion, our findings support the hypothesis of the acquisition of genomic regions from other pathogenic bacteria (*E. coli* or others) by horizontal transfers and reflect the genomic plasticity of EHEC or even *E. coli* strains. This variation in the genome contents of *E. coli*, suggested as a evolutionary strategy to better survive by Mokady *et al.* (2005), could lead to serious problems in public health and to the emergence of highly virulent new strains if one strain could acquire several strong virulence systems from different pathogenic bacteria, as it was dramatically illustrated by the 2011 Shiga toxin–producing *E. coli* O104:H4 German outbreak (Denamur, 2011; Rasko *et al.*, 2011).

Acknowledgements

During this study, Marjorie Bardiau was a PhD fellow of the 'Fonds pour la formation à la Recherche dans l'Industrie et dans l'Agriculture' (FRIA). This study was funded by the Federal Public Service of Health, Food Chain Safety and Environment (contract RF 6172), the European Network of Excellence EADGENE (European Animal Disease Genomics Network of Excellence for Animal Health and Food Safety) for the sequencing, and a grant 'Crédits aux chercheurs' FNRS (Fonds de Recherche Scientifique) 2008, no. 1363.

References

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA & Struhl K (1994) *Current Protocols in Molecular Microbiology*. Chapter 3. John Wiley and Sons, New York, NY.

Bardiau M, Labrozzo S & Mainil JG (2009) Putative adhesins of enteropathogenic and enterohemorrhagic *Escherichia coli* of serogroup O26 isolated from humans and cattle. *J Clin Microbiol* **47**: 2090–2096.

Bettelheim KA (2007) The non-O157 shiga-toxigenic (verocytotoxigenic) *Escherichia coli*; under-rated pathogens. *Crit Rev Microbiol* **33**: 67–87.

Brooks JT, Sowers EG, Wells JG, Greene KD, Griffin PM, Hoekstra RM & Strockbine NA (2005) Non-O157 Shiga toxin-producing *Escherichia coli* infections in the United States, 1983–2002. *J Infect Dis* **192**: 1422–1429.

Brzuszkiewicz E, Gottschalk G, Ron E, Hacker J & Dobrindt U (2009) Adaptation of pathogenic *E. coli* to various niches: genome flexibility is the key. *Genome Dyn* **6**: 110–125.

Caprioli A, Morabito S, Brugère H & Oswald E (2005) Enterohaemorrhagic *Escherichia coli*: emerging issues on virulence and modes of transmission. *Vet Res* **36**: 289–311.

China B, Pirson V & Mainil J (1996) Typing of bovine attaching and effacing *Escherichia coli* by multiplex *in vitro* amplification of virulence-associated genes. *Appl Environ Microbiol* **62**: 3462–3465.

Cobbaut K, Houf K, Buvens G, Habib I & De Zutter L (2009) Occurrence of non-sorbitol fermenting, verocytotoxin-lacking *Escherichia coli* O157 on cattle farms. *Vet Microbiol* **138**: 174–178.

Denamur E (2011) The 2011 Shiga toxin-producing *Escherichia coli* O104:H4 German outbreak: a lesson in genomic plasticity. *Clin Microbiol Infect* **17**: 1124–1125.

Durso LM, Bono JL & Keen JE (2005) Molecular serotyping of *Escherichia coli* O26:H11. *Appl Environ Microbiol* **71**: 4941–4944.

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.

Hornitzky MA, Mercieca K, Bettelheim KA & Djordjevic SP (2005) Bovine feces from animals with gastrointestinal infections are a source of serologically diverse atypical enteropathogenic *Escherichia coli* and Shiga toxin-producing *E. coli* strains that commonly possess intimin. *Appl Environ Microbiol* **71**: 3405–3412.

© 2012 Federation of European Microbiological Societies
Published by Blackwell Publishing Ltd. All rights reserved

Juhas M, van der Meer JR, Gaillard M, Harding RM, Hood DW & Crook DW (2009) Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol Rev* **33**: 376–393.

Kelly BG, Vespermann A & Bolton DJ (2009) Horizontal gene transfer of virulence determinants in selected bacterial foodborne pathogens. *Food Chem Toxicol* **47**: 969–977.

Kerr P, Ball H, China B, Mainil J, Finlay D, Pollock D, Wilson I & Mackie D (1999) Use of a monoclonal antibody against an *Escherichia coli* O26 surface protein for detection of enteropathogenic and enterohemorrhagic strains. *Clin Diagn Lab Immunol* **6**: 610–614.

Mainil JG, Jacquemin E, Herault F & Oswald E (1997) Presence of pap-, sfa-, and afa-related sequences in necrotoxigenic *Escherichia coli* isolates from cattle: evidence for new variants of the AFA family. *Can J Vet Res* **61**: 193–199.

Mokady D, Gophna U & Ron EZ (2005) Extensive gene diversity in septicemic *Escherichia coli* strains. *J Clin Microbiol* **43**: 66–73.

Moxley RA & Francis DH (1986) Natural and experimental infection with an attaching and effacing strain of *Escherichia coli* in calves. *Infect Immun* **53**: 339–346.

Nataro J & Kaper J (1998) Diarrheagenic *Escherichia coli*. *Clin Microbiol Rev* **11**: 142–201.

Ogura Y, Ooka T, Iguchi A *et al.* (2009) Comparative genomics reveal the mechanism of the parallel evolution of O157 and non-O157 enterohemorrhagic *Escherichia coli*. *P Natl Acad Sci USA* **106**: 17939–17944.

Ooka T, Terajima J, Kusumoto M *et al.* (2009) Development of a multiplex PCR-based rapid typing method for enterohemorrhagic *Escherichia coli* O157 strains. *J Clin Microbiol* **47**: 2888–2894.

Rasko DA, Webster DR, Sahl JW *et al.* (2011) Origins of the *E. coli* strain causing an outbreak of hemolytic-uremic syndrome in Germany. *N Engl J Med* **365**: 709–717.

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.

Stordeur P, China B, Charlier G, Roels S & Mainil J (2000) Clinical signs, reproduction of attaching/effacing lesions, and enterocyte invasion after oral inoculation of an O118 enterohaemorrhagic *Escherichia coli* in neonatal calves. *Microbes Infect* **2**: 17–24.

Szalo IM, Taminiau B, Goffaux F, Pirson V, McCappin J, Ball HJ & Mainil JG (2004) 2F3 monoclonal antibody recognizes the O26 O-antigen moiety of the lipopolysaccharide of enterohemorrhagic *Escherichia coli* strain 4276. *Clin Diagn Lab Immunol* **11**: 532–537.

Wells JG, Davis BR, Wachsmuth IK, Riley LW, Remis RS, Sokolow R & Morris GK (1983) Laboratory investigation of hemorrhagic colitis outbreaks associated with a rare *Escherichia coli* serotype. *J Clin Microbiol* **18**: 512–520.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. Dendrogram of 60 EPEC and EHEC strains constructed by PFGE data (UPGMA).

Table S1. Origin, pathotypes and serotypes of the *E. coli* strains of serogroup O26 tested for the distribution of specific sequences of the bovine EHEC strain 4276 (n.i., no information; w, week; m, month; y, year).

Table S2. Primers used in this study.

Table S3. BLASTN and BLASTX results of the 115 contigs specific to the bovine strain.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

© 2012 Federation of European Microbiological Societies
Published by Blackwell Publishing Ltd. All rights reserved