

## Multiple Insertional Events, Restricted by the Genetic Background, Have Led to Acquisition of Pathogenicity Island II<sub>J96</sub>-Like Domains among *Escherichia coli* Strains of Different Clinical Origins

Philippe Bidet, Stéphane Bonacorsi, Olivier Clermont, Caroline De Montille, Naima Brahim, and Edouard Bingen\*

Laboratoire d'études de génétique bactérienne dans les infections de l'enfant (EA3105),  
Université Denis Diderot—Paris 7, Service de Microbiologie,  
Hôpital Robert Debré (AP-HP), 75019 Paris, France

Received 16 November 2004/Returned for modification 3 January 2005/Accepted 9 March 2005

We investigated the dissemination of pathogenicity island (PAI) II<sub>J96</sub>-like elements (*hra*, *hly*, *cnf1*, and *pap*) among 455 *Escherichia coli* isolates from children and adults with urinary tract infection (UTI), neonates with meningitis or colonized healthy neonates, and 74 reference strains by means of PCR phylogenetic grouping, ribotyping, and PCR analysis of virulence genes. Colocalization of these genes was documented by pulsed-field gel electrophoresis followed by Southern hybridization and long-range PCR (LRPCR) between the *hra* and the *papG* alleles. Site-specific insertion of the PAI was determined by LRPCR between *hra* and tRNA flanking sequences. *hra*, *hly*, and *cnf1* were found in 113 isolates and consistently colocalized, constituting the backbone of PAI II<sub>J96</sub>-like domains. The prevalence of PAI II<sub>J96</sub>-like domains was significantly higher among UTI isolates than among neonatal meningitis and commensal isolates. These domains were restricted to a few ribotypes of group B2. In contrast to the consistent colocalization of *hra*, *hly*, and *cnf1*, the *pap* operon was varied: 12% of strains exhibited an allelic exchange of the *papG* class III allele (*papGIII*) for the *papG* class II allele (*papGII*) (only UTI isolates), and the *pap* operon was deleted in 23% of strains. No strains harbored *papGIII* outside the PAI, which appears to be the only source of this allele. PAI II<sub>J96</sub>-like domains were inserted in the vicinities of three different tRNAs—*pheU* (54%), *leuX* (29%), and *pheV* (15%)—depending on the genetic backgrounds and origins of the isolates. Multiple insertional events restricted by the genetic background have thus led to PAI II<sub>J96</sub> acquisition. Specific genetic backgrounds and insertion sites may have played a role in additional recombination processes for *E. coli* adaptation to different ecological niches.

*Escherichia coli* is a normal inhabitant of the human intestinal tract but is also a leading cause of community-acquired infections. In addition to causing intestinal infections, *E. coli* is the most frequent cause of gram-negative bacterial infections such as cystitis, pyelonephritis, bacteremia, and neonatal meningitis. These extraintestinal pathogenic *E. coli* (ExPEC) strains (38) differ from commensal *E. coli* strains in two major respects. First, among the four main phylogenetic groups of *E. coli* (A, B1, B2, and D), ExPEC strains belong mostly to group B2 and, to a lesser extent, to group D, whereas commensal strains belong mainly to group A (2, 8, 25, 36). Second, ExPEC strains harbor many genetic virulence determinants and other fitness factors. Most of these genes are acquired by horizontal transfer and constitute the so-called “ectochromosomal” DNA or “flexible gene pool.” The pathogenicity island (PAI) is one of the most important elements of ectochromosomal DNA (17). These large chromosomal regions (>10 kb), differing in their G+C contents from that of the core genome, are located near tRNA genes and contain both genetic virulence determinants and mobility genes (19).

Although the relationship between phylogenetic groups and extraintestinal virulence genes is well documented, few studies have focused on the relationship between the PAI and the genetic background of recipient *E. coli* isolates from different clinical sources (14, 36).

The aim of this study was to examine the interaction between the genetic background of *E. coli* strains and the integration and evolution of PAIs according to the ecological niche. Among the archetypal PAIs described for ExPEC, PAI II<sub>J96</sub> appeared to be a good candidate for such an investigation. PAI II<sub>J96</sub>, initially described to occur in the uropathogenic *E. coli* strain J96, is one of the largest PAIs described to date (~110 kb) (7, 23). It contains at least four genes or operons, including those coding for hemolysin (*hly*), cytotoxic necrotizing factor (*cnf1*), P fimbriae (*pap*) with the variant allelic adhesin (*papG* class III allele [*papGIII*]), and heat-resistant agglutinin (*hra*) (7, 40). PAI II<sub>J96</sub> contributes to the virulence of cystitis, pyelonephritis, and neonatal meningitis strains (8, 15, 19, 22, 30). A previous study of urosepsis isolates suggests that three colocalized genes—*hly*, *cnf1*, and *hra*—constitute the backbone of a PAI II<sub>J96</sub>-like domain (6). Here we examined the distribution and insertion sites of PAI II<sub>J96</sub>-like domains in a large collection of ExPEC isolates from various clinical settings with regard to their phylogenetic groups and subgroups.

\* Corresponding author. Mailing address: Service de Microbiologie, Hôpital Robert Debré, 48 Bd. Sérurier, 75395 Paris cedex 19, France. Phone: 33 1 40 03 23 40. Fax: 33 1 40 03 24 50. E-mail: edouard.bingen@rdb.ap-hop-paris.fr.

TABLE 1. Oligonucleotide primers used to amplify virulence-associated genes and for long-range PCR

Primer designation	Primer sequence	Target <sup>a</sup>	Size of PCR product (bp)	Reference of source
chuA.1 chuA.2	5'-GACGAACCAACGGTCAGGAT-3' 5'-TGCCGCCAGTACCAAAGACA-3'	<i>chuA</i>	279	11
TspE4C2.1 TspE4C2.2	5'-GAGTAATGTCGGGGCATTCA-3' 5'-CGCGCCAACAAAGTATTACG-3'	<i>tspE4.C2</i>	152	11
yjaA.1 yjaA.2	5'-TGAAGTGTGAGGAGACGCTG-3' 5'-ATGGAGAATGCGTTCCTCAAC-3'	<i>yjaA</i>	211	11
hly.1 hly.2	5'-AGGTTCTTGGGCATGTATCCT-3' 5'-TTGCTTTGCAGACTGCAGTGT-3'	<i>hlyC/A</i>	556	2
cnfI.1 cnfI.2	5'-CAGTGACCGGATCTCCGTTAT-3' 5'-CGTGTAATTCTTCTGTACTTCC-3'	<i>cnfI</i>	230	35
hra.1 hra.2	5'-CAGAAAAACAACCGGTATCAG-3' 5'-ACCAAGCATGATGTCATGAC-3'	<i>hra</i>	260	6
papC.1 papC.2	5'-GACGGCTGTACTGCAGGGTGTGGCG-3' 5'-ATATCCTTTCTGCAGGGATGCAATA-3'	<i>papC</i>	328	33
papG.II.1 papG.II.2	5'-GGGATGAGCGGGCCTTTGAT-3' 5'-CGGGCCCCCAAGTAACTCG-3'	<i>papGII</i>	190	24
papG.III.1 papG.III.2	5'-GGCCTGCAATGGATTTACCTGG-3' 5'-CCACCAAATGACCATGCCAGAC-3'	<i>papGIII</i>	258	24
pheR.1 pheR.2	5'-GCCGCAATCTTAAGCAGTTG-3' 5'-GCACGACATTTACGTCAGT-3'	<i>pheU</i>	350	6
pheR.1 yjgB.1	5'-GCCGCAATCTTAAGCAGTTG-3' 5'-ACCTTGCTCGCAGTTGATCT-3'	<i>pheU</i> (LR PCR) <i>leuX</i> ( <i>yjgB</i> ) (LR PCR)		6 6
PheV.1	5'-AACCGGATTACGCATCTGTG-3'	<i>pheV</i> (LR PCR)		This study

<sup>a</sup> LR PCR, long-range PCR with hra.1. C2, anonymous DNA fragment issued from subtractive hybridization (11).

## MATERIALS AND METHODS

**Bacterial strains.** We analyzed 455 clinical *E. coli* isolates recovered from 1997 to 2000. They consisted of a previously described series of 100 French adult urosepsis isolates (6), 134 international strains of *E. coli* causing neonatal meningitis (ECNM) (8), and 84 urinary tract infection (UTI) isolates from French infants (<90 days) (9), as well as unpublished French collections of 75 UTI isolates from children aged from 3 months to 10 years and 62 isolates colonizing healthy neonates. All isolates were stored at -80°C until characterization.

Reference strains carrying a PAI II<sub>96</sub>-like domain, uropathogenic *E. coli* (UPEC) strains AD110 and J96, and the 72 strains of the ECOR collection were studied for comparison (23, 34, 41).

**Phylogenetic grouping and subgrouping.** The main phylogenetic groups (A, B1, B2, and D) were determined for all strains by using previously described PCR methods (11), and B2 strains were subgrouped by ribotyping with the restriction enzyme HindIII and with 16S and 23S rRNAs as the probes (1, 3-5).

**Detection of virulence genes and characterization of PAI II<sub>96</sub>-like domains.** Each strain was screened for *hlyC*, *hlyA*, *cnfI*, *papC*, *papG* class II and class III alleles, and *hra* by means of PCR, as previously described (Table 1) (6, 24, 33, 35), and their colocalizations were detected by pulsed-field gel electrophoresis using the restriction enzyme NotI followed by Southern hybridization (6). When strains harbored the two *papG* alleles, long-range PCR between the forward *papG* allele primers and the reverse *hra* primer hra.1 was used to determine which *papG* allele belonged to the PAI II<sub>96</sub>-like domain; the Expand Long Template PCR system (Roche) was used as previously described (6). The insertion sites of PAI II<sub>96</sub>-like domains were determined by using the same long-range PCR method between *hra* (primer hra.1) and either *pheU* (formerly *pheR*) or *leuX* tRNA flanking sequences, as previously described (Table 1) (6). When these PCRs were negative, isolates were screened for an insertion in *pheV* tRNA by long-range PCR with primers homologous to *hra* (primer hra.1) and the *pheV* tRNA flanking sequence (primer PheV.1) (Table 1). This tRNA was chosen because it has been described as a second insertion site for a PAI also inserted in *pheU* (31). To control the integrity of the archetypal *pheU* tRNA insertion site,

we performed PCR of the flanking sequences of the archetypal insertion site *pheU* with primers pheR.1 and pheR.2 (Table 1).

**Statistical analysis.** Fisher's exact test was used. *P* values of <0.05 were considered statistically significant.

## RESULTS

Among the 455 clinical isolates and the reference strains, *hly*, *hra*, and *cnfI* were present simultaneously, and always colocalized, in 113 strains (104 of the 455 clinical isolates [23%], 7 of the 72 ECOR strains [10%], and the UPEC reference strains J96 and AD110). None of our 455 clinical isolates carried *cnfI* without *hly* and *hra*. The distribution of these strains is shown in Table 2 according to their clinical sources. PAI II<sub>96</sub>-like domains were significantly more frequent among UTI isolates (adult urosepsis and children and infant UTI; 30% of 259 isolates) than among ECNM and neonatal colonization isolates (10% of 134 isolates and 19% of 62 isolates, respectively) (*P* < 0.01). In our ECNM collection, PAI II<sub>96</sub>-like domains were mostly found among the 38 O18:K1 strains (in 9 strains, or 24%).

The distribution of these PAI II<sub>96</sub>-like domains was restricted to a limited number of genetic backgrounds. Indeed, our clinical collection comprised 46, 9, 305, and 95 isolates from phylogenetic groups A, B1, B2, and D, respectively (Table 2), while all 113 strains harboring a PAI II<sub>96</sub>-like domain belonged only to group B2, apart from two group D strains.

TABLE 2. Presence of a PAI II<sub>J96</sub>-like domain according to the phylogenetic group and subgroup among collections of *E. coli* strains of various origins

Phylogenetic group/ subgroup <sup>a</sup>	Adults with urosepsis ( <i>n</i> = 100)		Children with UTI ( <i>n</i> = 75)		Infants with UTI ( <i>n</i> = 84)		Neonates with meningitis ( <i>n</i> = 134)		Neonates with colonization ( <i>n</i> = 62)		ECOR collection ( <i>n</i> = 72)	
	No.	PAI II <sub>J96</sub> <sup>b</sup>	No. (%) <sup>c</sup>	PAI II <sub>J96</sub>	No. (%)	PAI II <sub>J96</sub>	No. (%)	PAI II <sub>J96</sub>	No. (%)	PAI II <sub>J96</sub>	No. (%)	PAI II <sub>J96</sub>
A ( <i>n</i> = 71)	11	0	3 (4)	0	7 (8)	0	12 (9)	0	13 (21)	0	25 (35)	0
B1 ( <i>n</i> = 25)	1	0	1	0	2	0	3 (2)	0	2	0	16 (22)	0
B2 ( <i>n</i> = 320)	61	21 (34)	51 (68)	32 (63)	60 (71)	25 (42)	96 (72)	14 (15)	37 (60)	11 (30)	15 (21)	6 (40)
Ribotype I	20	13 (65)	ND	19	16 (19)	12 (75)	4 (3)	2	1	1	6 (8)	3 (50)
Ribotype II	19	0	ND	0	23 (27)	1	55 (41)	9 (16)	18 (29)	1	2	0
Ribotype III	12	3 (25)	ND	8	15 (18)	9 (60)	15 (11)	3 (20)	8 (13)	3 (38)	2	1
Ribotype IX	4	3 (75)	ND	3	2	2	6 (5)	0	6 (10)	6 (100)	0	0
Ribotype X	4	0	ND	0	2	0	2	0	4 (7)	0	1	0
Ribotype XI	2	2	ND	2	1	1	0	0	0	0	3 (4)	2
Other ribotypes	0	0	ND	0	1	0	14 (10)	0	0	0	1	0
D ( <i>n</i> = 107)	27	0	20 (27)	0	15 (18)	0	23 (17)	0	10 (16)	1	12 (17)	1
Other ( <i>n</i> = 4)	0	0	0	0	0	0	0	0	0	0	4 (5)	0
Total	100	21 (21)	75	32 (43)	84	25 (30)	134	14 (10)	62	12 (19)	72	7 (10)

<sup>a</sup> Each subgroup corresponds to a different ribotype, as described previously (6).

<sup>b</sup> Number (and percentage) of strains harboring a PAI II<sub>J96</sub>-like domain among each phylogenetic group or subgroup of the collection.

<sup>c</sup> ND, not determined for the entire collection.

Twelve ribotypes were identified among the 305 group B2 clinical isolates, whereas strains harboring a PAI II<sub>J96</sub>-like domain belonged to only five ribotypes (Table 3).

In contrast to the consistent association of *hra*, *hly*, and *cnf1*, the *pap* operon was more varied: 23% of *hra*-, *hly*-, and *cnf1*-positive strains did not harbor a colocalized *pap* operon, and 12% bore the *papGII* allele instead of the *papGIII* allele within the PAI, as shown by positive *papGII-hra* long-range PCR (Table 4). The 23% of *pap*-negative strains harboring a PAI II<sub>J96</sub>-like domain included six ECNM isolates (43%), three urosepsis isolates (14%), five childhood UTI isolates (20%), nine infantile UTI isolates (28%), and two colonization isolates (16%). Three different PAI II<sub>J96</sub>-like domains were defined on the basis of *pap* operon variations, namely, the *papGIII*-positive, *papGII*-positive, and *pap*-negative domains. The *papGIII* allele was found in none of the 351 isolates negative for the PAI II<sub>J96</sub>-like domain, and no isolates harbored a *papGIII* allele outside the PAI II<sub>J96</sub>-like domain. In contrast, *papGII* was found outside the PAI in 56% of UTI isolates versus in 7% and 8% of ECNM and colonization isolates, respectively ( $P < 0.01$ ). When located outside the PAI II<sub>J96</sub>-like domain, the *papGII* allele was physically linked to *hly* in seven strains.

The PAI II<sub>J96</sub>-like domains were inserted in three different tRNAs, namely, *pheU* (formerly *pheR*) (54%), *leuX* (29%), and *pheV* (15%). The insertion site was unknown for two isolates (Table 4). When a PAI II<sub>J96</sub>-like domain was inserted in *pheU*, *pheU* PCR was negative owing to tRNA disruption. When PAI II<sub>J96</sub>-like domains were inserted in *leuX* or *pheV*, *pheU* PCR was positive in 64% and 94% of cases, respectively, showing that the *pheU* site was free. The insertion site differed according to the genetic background determined by ribotyping. In ribotype I and ribotype XI isolates, the insertion occurred mostly in the *pheU* tRNA (71% and 100%, respectively, versus 35% for other ribotypes;  $P < 0.01$ ), whereas it occurred mainly in *pheV* in ribotype IX isolates (Table 3). When the insertion

occurred in *leuX*, long-range PCR between *hra* and *leuX* yielded a 4.5-kbp product for all ribotype IX isolates and an 8-kbp product for isolates of other ribotypes. When the insertion occurred in *pheU*, long-range PCR between *hra* and *pheU* yielded a 5.2-kbp product for all isolates belonging to ribotypes II, IX, and XI and a product of either 5.2 kbp or 7.5 kbp for isolates belonging to ribotypes I and III.

The frequencies of *papGIII* differed according to the insertion site. *papGIII* was significantly more frequent in PAI II<sub>J96</sub>-like domains inserted in *leuX* than in PAI II<sub>J96</sub>-like domains inserted in *pheU* or *pheV* (94% versus 58% or 47%, respectively;  $P < 0.01$ ) (Table 4).

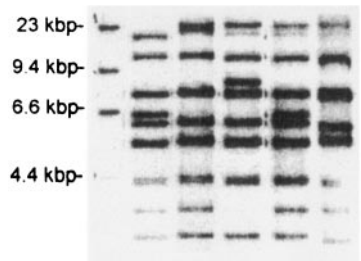
The replacement of *papGIII* by *papGII* observed in PAI II<sub>J96</sub>-like domains inserted in *pheU* (15%) and, to a lesser extent, in *leuX* (two isolates) was not observed in domains inserted in *pheV* (Table 4).

## DISCUSSION

The consistent colocalization of *hly*, *hra*, and *cnf1* in 113 of 529 *E. coli* strains of various origins, contrasting with the marked plasticity of the *pap* operon, considerably extends our previous results for adult urosepsis isolates, in which the *hly*, *cnf1*, and *hra* gene triplet constitutes the backbone of the PAI II<sub>J96</sub>-like domain (6). The presence of these specific genes does not necessarily imply the presence of the complete PAI II<sub>J96</sub>, as we did not analyze the flanking sequences. Therefore, for a given strain, the simultaneous detection of these three genes may be considered the signature of this ectochromosomal domain. Although the consistent association of *hly* and *cnf1* could be explained by a combined cytotoxic effect, the role of *hra* in this group of genes remains to be determined (15, 32). Interestingly, in the overall strain collection studied here, the *papGIII* allele was found in only 64% of strains harboring this backbone and was consistently colocalized. Other authors studying both *papGIII* and *cnf1* have never found isolates harboring *papGIII*

TABLE 3. Molecular characterization of the PAI II<sub>196</sub>-like domains and clinical origins of group B2 strains according to ribotype

Ribotype <sup>c</sup> (no. of isolates)	No. (%) of isolates										Reference strain			
	Insertion site <sup>a</sup>					Clinical origin <sup>b</sup>								
	<i>papG</i> allele <sup>a</sup>		Adults with urosepsis (n = 21)			Children with UTI (n = 32)		Infants with UTI (n = 25)	Neonates with meningitis (n = 14)	Neonates with colonization (n = 11 <sup>d</sup> )		ECOR (n = 6 <sup>e</sup> )		
	<i>pheU</i>	<i>leuX</i>	<i>pheV</i>	Unknown	<i>papGIII</i>	<i>papGII</i>	Absent							
I (51)	35 (71)	9 (18)	5 (10)	2	30 (59)	6 (12)	15 (29)	13 (62 <sup>e</sup> )	19 (59 <sup>e</sup> )	12 (48 <sup>e</sup> )	2	1	3	AD110
II (11)	4 (36)	6 (54)	1		6 (54.5)		5 (45.5)			1	9 (64 <sup>f</sup> )	1		
III (27)	13 (48)	11 (41)	3 (11)		15 (56)	6 (22)	6 (22)	3 (14)	8 (25)	9 (36)	3 (21)	3 (25)	1	
IX (14)	2 <sup>g</sup>	6 <sup>g</sup> (43)	7 (50 <sup>h</sup> )		14 (100)			3 (14)	3 (9)	2		6 (50 <sup>i</sup> )		
XI (8)	8 (100)				7 (87.5)	1		2	2	1			2	J96



<sup>a</sup> Numbers in parentheses are percentages of strains among each ribotype.  
<sup>b</sup> Numbers in parentheses are percentages of strains among each clinical collection.  
<sup>c</sup> The number of the ribotype corresponds to the ribotyping patterns described previously (6).  
<sup>d</sup> One strain of this collection belonged to group D.  
<sup>e</sup> By the Fisher test, *P* was <0.05 in a comparison with values for meningitis and colonization isolates.  
<sup>f</sup> By the Fisher test, *P* was <0.05 in a comparison with values for UTI and colonization isolates.  
<sup>g</sup> One isolate of ribotype IX harbored two PAI II<sub>196</sub>-like domains inserted in *pheU* and *leuX*.  
<sup>h</sup> By the Fisher test, *P* was <0.05 in a comparison with values for the other ribotypes.  
<sup>i</sup> By the Fisher test, *P* was <0.05 in a comparison with values for meningitis and UTI isolates.

TABLE 4. Presence of *papG* alleles and origin of the strains according to the tRNA insertion sites

Location of tRNA insertion site (no. of isolates)	No. (%) of isolates								ECOR (n = 7)
	<i>papG</i> allele <sup>a</sup>			Clinical origin <sup>b</sup>					
	Class III (65%)	Class II (12%)	Absent (23%)	Adults with urosepsis (n = 21 <sup>c</sup> )	Children with UTI (n = 32)	Infants with UTI (n = 25)	Neonates with meningitis (n = 14)	Neonates with colonization (n = 12)	
<i>pheU</i> (62 <sup>d</sup> )	36 <sup>d</sup> (58)	9 (15)	17 (27)	18 (86)	21 (66)	9 (36)	7 (50)	2	3
<i>leuX</i> (33)	31 (94)	2		4 (19)	5 (15)	12 (48)	7 (50)	3 (25)	2
<i>pheV</i> (17)	8 (47)		9 (53)		6 (19)		2	7 (58 <sup>e</sup> )	2
Unknown (2)		2					2		

<sup>a</sup> Numbers in parentheses are percentages of strains among each insertion site.

<sup>b</sup> Numbers in parentheses are percentages of strains among each clinical collection.

<sup>c</sup> One isolate of ribotype IX harbored two PAI II<sub>J96</sub>-like domains inserted in *pheU* and *leuX*.

<sup>d</sup> The two UPEC reference strains J96 and AD110 have their PAI II<sub>J96</sub> inserted in *pheU* with a *papGIII* allele.

<sup>e</sup> By the Fisher test, *P* was <0.05 in a comparison with values for meningitis and UTI isolates.

without *cnfI* (28, 29). This suggests that archetypal PAI II<sub>J96</sub> may be the sole source of the *papGIII* allele and that it has evolved by the allele substitution of *papG* or by the deletion of the *pap* operon (6). In contrast, the *papGII* allele was found instead of *papGIII* within and/or outside the PAI (data not shown). The colocalization of *papGII* with *hly* outside the PAI II<sub>J96</sub>-like domain in seven strains suggests that *papGII* is located in another PAI. Deletion may optimize the structure of PAI elements and reduce the genetic burden by eliminating genes whose products are no longer used (17). However, *E. coli* strains that lack virulence factors are able to cause extraintestinal infections, including UTI, in compromised hosts (6, 26, 28).

As previously reported, PAI II<sub>J96</sub>-like domains were almost exclusively restricted to group B2 (6); the only exceptions were two group D strains. Furthermore, among group B2 strains, PAI II<sub>J96</sub>-like domains were restricted to only 5 of the 12 ribotypes identified among all strains studied. This points to a strong association between PAI II<sub>J96</sub>-like domains and a few B2 genetic backgrounds. Two scenarios of PAI II<sub>J96</sub> acquisition may explain these data. In the first scenario, PAI II<sub>J96</sub> was acquired once, by chance, by a common B2 ancestor of these five subgroups and was subsequently transmitted vertically and eventually rearranged and deleted with additional recombination processes leading to additions or deletions within the PAIs (19, 27, 40). In the second scenario, the integration of this PAI in a group B2 *E. coli* strain occurred by multiple insertional events which were restricted to these five genetic backgrounds because they are compatible with PAI integration and expression (19). Other studies have also suggested that specific genetic backgrounds are required for the integration, retention, and expression of PAIs acquired by several horizontal transfers among ExPEC strains (14, 28). Acquired sequences are effective only if their expression is coordinated with that of the rest of the chromosome and with the life cycle of the microbial host (16). The second scenario appears more likely, because we found that PAI II<sub>J96</sub>-like domains were inserted within at least three different tRNAs. To our knowledge, this is the first description of three different insertion sites for a given PAI-like domain in *E. coli*. Only the high-pathogenicity island in *Yersinia* spp. displays such a distribution of insertion sites (10, 39). However, a multiplicity of high-pathogenicity island insertion

sites can occur in a single strain of *Yersinia* spp., due to the sequence identity of the different *asn* tRNAs; this is not the case for *pheU* and *leuX* in PAI II<sub>J96</sub>-like domains.

Virulence factors carried by the genetic background may influence PAI acquisition. Indeed, in our clinical isolates, the *papGIII*-positive PAI II<sub>J96</sub>-like domain was always associated with *papGII* outside the PAI in isolates of ribotype I, while the *papGIII*-positive PAI II<sub>J96</sub>-like domain, with or without *papGII*, was present in isolates belonging to the other four ribotypes (data not shown). The fact that *papGII* alone (without the PAI II<sub>J96</sub>-like domain) was present in isolates of ribotype I, contrary to what occurs with *papGIII*, suggests that *papGII* was acquired first by isolates of ribotype I. So, it is tempting to speculate that the PAI II<sub>J96</sub>-like domain was acquired secondarily, only on a ribotype I background still carrying *papGII* or genetic determinants belonging to PAIs containing *papGII*. Of note, we found no isolates harboring *papC* without a *papGII* or *papGIII* allele within the PAI II<sub>J96</sub>-like domain. Thus, the *papGI* allele was not present in the PAI II<sub>J96</sub>-like domains of our collection.

The genetic background also seems to influence the insertion site of the mobile genetic element carrying the PAI. Indeed, both the frequencies of site-specific PAI insertion in the different tRNAs and the lengths of the PCR product between *hra* and the tRNA differed with the ribotype. To examine whether insertion in a given tRNA occurs randomly or is restricted by free insertion sites, we performed PCR of the flanking sequences of the archetypal insertion site *pheU*. When PAI II<sub>J96</sub>-like domains were inserted in *leuX* or *pheV*, *pheU* PCR was positive in 64% or 94% of cases, respectively, showing that the *pheU* site was free. These results support the possible influence of the genetic background on the insertion site.

The key features of this PAI that render it incompatible with many other genetic backgrounds remain to be determined. *cnfI*, which has never been described to occur in another ectochromosomal DNA, may offer one line of investigation (32).

Of particular interest was the influence of the insertion site on the plasticity of the PAI. Indeed, the PAI II<sub>J96</sub>-like domains inserted in *leuX* displayed lower degrees of variation than those inserted in other sites, as the *pap* operon with the *papGIII* allele was present in 94% of cases, compared to 58% in *pheU* tRNA and 47% in *pheV* tRNA (*P* < 0.01). Dobrindt et



al. found that, in *E. coli* 536, *leuX* was required for the efficient expression of several virulence genes, such as *hly* and type 1 fimbriae (13, 37). Our results show that PAI plasticity is dependent on the insertion site, whatever the B2 subgroup or the clinical origin. Interestingly, the allelic exchange of *papGII* for *papGIII*, observed mostly in PAI II<sub>J96</sub>-like domains inserted in *pheU*, was not observed in *pheV*. This difference in *papG* alleles may influence pathogenicity, particularly as it was encountered only among UTI isolates (18% versus 0% of ECNM and colonization isolates;  $P < 0.01$ ). Moreover, 56% of UTI isolates harbored a chromosomal *papGII* allele outside the PAI II<sub>J96</sub>-like domain. The allelic switch to *papGII* within the PAI II<sub>J96</sub>-like domain, and/or the acquisition of *papGII* outside this PAI, may allow the bacterium to colonize the urinary tract.

When we examined the distribution of genetic backgrounds among clinical isolates of different sources carrying PAI II<sub>J96</sub>-like domains, we found that ribotype II predominated among ECNM isolates and ribotype I predominated among UTI isolates (Table 3). However, this distribution also reflects the predominance of these ribotypes in the overall strain collections (Table 2). Houdouin et al. found that, in ribotype II strains, this PAI contributed to bacterial survival in blood by inducing high-level bacteremia, a step preceding blood-brain barrier penetration (22). Thus, ECNM isolates of ribotype II lacking PAI II<sub>J96</sub> may carry other virulence factors. The predominance of ribotype IX in neonatal colonization isolates carrying a PAI II<sub>J96</sub>-like domain contrasted with the significantly lower prevalence of this ribotype in the entire collection of neonatal colonization isolates (Table 2). Moreover, all but one of the colonization isolates of ribotype IX had their PAI II<sub>J96</sub>-like domains inserted in *pheV*, with a *papGIII*-positive *pap* operon. This raises the possibility that PAI insertion in *pheV* on a particular genetic background may alter virulence gene expression or favor the expression of fitness factors contributing to increasing survival in the gut, thus forming a “saprophytic island” (12, 20, 21).

In conclusion, the simultaneous detection of *hly*, *cnf1*, and *hra* may be considered the signature of a PAI II<sub>J96</sub>-like domain in a given strain of *E. coli* and could be used for further epidemiological studies. Multiple insertional events at at least three different sites, restricted by the genetic background, have thus led to PAI II<sub>J96</sub>-like domain acquisition. Specific genetic backgrounds and insertion sites may have played a role in additional recombination processes for *E. coli* adaptation to different ecological niches (18, 20).

#### REFERENCES

- Bingen, E., E. Denamur, N. Lambert-Zechovsky, Y. Aujard, N. Brahimi, P. Geslin, and J. Elion. 1992. Analysis of DNA restriction fragment length polymorphism extends the evidence for breast milk transmission in *Streptococcus agalactiae* late-onset neonatal infection. *J. Infect. Dis.* **165**:569–573.
- Bingen, E., B. Picard, N. Brahimi, S. Mathy, P. Desjardins, J. Elion, and E. Denamur. 1998. Phylogenetic analysis of *Escherichia coli* strains causing neonatal meningitis suggests horizontal gene transfer from a predominant pool of highly virulent B2 group strains. *J. Infect. Dis.* **177**:642–650.
- Bingen, E. H., E. Denamur, and J. Elion. 1994. Use of ribotyping in epidemiological surveillance of nosocomial outbreaks. *Clin. Microbiol. Rev.* **7**:311–327.
- Bingen, E. H., E. Denamur, N. Y. Lambert-Zechovsky, A. Bourdois, P. Mariani-Kurkdjian, J.-P. Cezard, J. Navarro, and J. Elion. 1991. DNA restriction fragment length polymorphism differentiates crossed from independent infections in nosocomial *Xanthomonas maltophilia* bacteremia. *J. Clin. Microbiol.* **29**:1348–1350.
- Bingen, E. H., E. Denamur, B. Picard, P. Goulet, N. Y. Lambert-Zechovsky, N. Brahimi, J.-C. Mercier, F. Beaufilets, and J. Elion. 1992. Molecular epidemiology unravels the complexity of neonatal *Escherichia coli* acquisition in twins. *J. Clin. Microbiol.* **30**:1896–1898.
- Bingen-Bidois, M., O. Clermont, S. Bonacorsi, M. Terki, N. Brahimi, C. Loukil, D. Barraud, and E. Bingen. 2002. Phylogenetic analysis and prevalence of urosepsis strains of *Escherichia coli* bearing pathogenicity island-like domains. *Infect. Immun.* **70**:3216–3226.
- Blum, G., V. Falbo, A. Caprioli, and J. Hacker. 1995. Gene clusters encoding the cytotoxic necrotizing factor type 1, Prs-fimbriae and alpha-hemolysin form the pathogenicity island II of the uropathogenic *Escherichia coli* strain J96. *FEMS Microbiol. Lett.* **126**:189–195.
- Bonacorsi, S., O. Clermont, V. Houdouin, C. Cordevant, N. Brahimi, A. Marecat, C. Tinsley, X. Nassif, M. Lange, and E. Bingen. 2003. Molecular analysis and experimental virulence of French and North American *Escherichia coli* neonatal meningitis isolates: identification of a new virulent clone. *J. Infect. Dis.* **187**:1895–1906.
- Bonacorsi, S., S. Lefevre, O. Clermont, V. Houdouin, A. Bourrillon, C. Loirat, Y. Aujard, and E. Bingen. 2005. *Escherichia coli* strains causing urinary tract infection in uncircumcised infants resemble urosepsis-like adult strains. *J. Urol.* **173**:195–197.
- Buchrieser, C., R. Brosch, S. Bach, A. Guiry, and E. Carniel. 1998. The high-pathogenicity island of *Yersinia pseudotuberculosis* can be inserted into any of the three chromosomal *asn* tRNA genes. *Mol. Microbiol.* **30**:965–978.
- Clermont, O., S. Bonacorsi, and E. Bingen. 2000. Rapid and simple determination of the *Escherichia coli* phylogenetic group. *Appl. Environ. Microbiol.* **66**:4555–4558.
- Connell, H., M. Hedlund, W. Agace, and C. Svanborg. 1997. Bacterial attachment to uro-epithelial cells: mechanisms and consequences. *Adv. Dent. Res.* **11**:50–58.
- Dobrindt, U., L. Emödy, I. Gentschev, W. Goebel, and J. Hacker. 2002. Efficient expression of the alpha-haemolysin determinant in the uropathogenic *Escherichia coli* strain 536 requires the *leuX*-encoded tRNA<sub>3<sup>Leu</sup></sub>. *Mol. Genet. Genomics* **267**:370–379.
- Escobar-Paramo, P., O. Clermont, A. B. Blanc-Potard, H. Bui, C. Le Bouguenec, and E. Denamur. 2004. A specific genetic background is required for acquisition and expression of virulence factors in *Escherichia coli*. *Mol. Biol. Evol.* **21**:1085–1094.
- Falbo, V., M. Famiglietti, and A. Caprioli. 1992. Gene block encoding production of cytotoxic necrotizing factor 1 and hemolysin in *Escherichia coli* isolates from extraintestinal infections. *Infect. Immun.* **60**:2182–2187.
- Groisman, E. A., and H. Ochman. 1997. How *Salmonella* became a pathogen. *Trends Microbiol.* **5**:343–349.
- Hacker, J., L. Bender, M. Ott, J. Wingender, B. Lund, R. Marre, and W. Goebel. 1990. Deletions of chromosomal regions coding for fimbriae and hemolysins occur in vitro and in vivo in various extraintestinal *Escherichia coli* isolates. *Microb. Pathog.* **8**:213–225.
- Hacker, J., G. Blum-Oehler, I. Muhldorfer, and H. Tschape. 1997. Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. *Mol. Microbiol.* **23**:1089–1097.
- Hacker, J., G. Blum-Oehler, B. Janke, G. Nagy, and W. Goebel. 1999. Pathogenicity islands of extraintestinal *Escherichia coli*, p. 59–76. In J. B. Kaper and J. Hacker (ed.), *Pathogenicity islands and other mobile virulence elements*. ASM Press, Washington, D.C.
- Hacker, J., and E. Carniel. 2001. Ecological fitness, genomic islands and bacterial pathogenicity. A Darwinian view of the evolution of microbes. *EMBO Rep.* **2**:376–381.
- Hacker, J., U. Hentschel, and U. Dobrindt. 2003. Prokaryotic chromosomes and disease. *Science* **301**:790–793.
- Houdouin, V., S. Bonacorsi, N. Brahimi, O. Clermont, X. Nassif, and E. Bingen. 2002. A uropathogenicity island contributes to the pathogenicity of *Escherichia coli* strains that cause neonatal meningitis. *Infect. Immun.* **70**:5865–5869.
- Hull, R. A., R. E. Gill, P. Hsu, B. H. Minshew, and S. Falkow. 1981. Construction and expression of recombinant plasmids encoding type 1 or D-mannose-resistant pili from a urinary tract infection *Escherichia coli* isolate. *Infect. Immun.* **33**:933–938.
- Johnson, J. R., and J. J. Brown. 1996. A novel multiply primed polymerase chain reaction assay for identification of variant *papG* genes encoding the Gal(alpha 1–4)Gal-binding *PapG* adhesins of *Escherichia coli*. *J. Infect. Dis.* **173**:920–926.
- Johnson, J. R., P. Delavari, M. Kuskowski, and A. L. Stell. 2001. Phylogenetic distribution of extraintestinal virulence-associated traits in *Escherichia coli*. *J. Infect. Dis.* **183**:78–88.
- Johnson, J. R., M. A. Kuskowski, T. T. O'Bryan, and J. N. Maslow. 2002. Epidemiological correlates of virulence genotype and phylogenetic background among *Escherichia coli* blood isolates from adults with diverse-source bacteremia. *J. Infect. Dis.* **185**:1439–1447.
- Johnson, J. R., T. T. O'Bryan, M. Kuskowski, and J. N. Maslow. 2001. Ongoing horizontal and vertical transmission of virulence genes and *papA* alleles among *Escherichia coli* blood isolates from patients with diverse-source bacteremia. *Infect. Immun.* **69**:5363–5374.
- Johnson, J. R., T. A. Russo, P. I. Tarr, U. Carlino, S. S. Bilge, J. C. Vary, Jr., and A. L. Stell. 2000. Molecular epidemiological and phylogenetic associa-

- tions of two novel putative virulence genes, *iha* and *iroN*<sub>*E. coli*</sub>, among *Escherichia coli* isolates from patients with urosepsis. *Infect. Immun.* **68**:3040–3047.
29. **Johnson, J. R., and A. L. Stell.** 2000. Extended virulence genotypes of *Escherichia coli* strains from patients with urosepsis in relation to phylogeny and host compromise. *J. Infect. Dis.* **181**:261–272.
  30. **Kim, K. S.** 2003. Pathogenesis of bacterial meningitis: from bacteraemia to neuronal injury. *Nat. Rev. Neurosci.* **4**:376–385.
  31. **Lalioui, L., and C. Le Bouguéneq.** 2001. *afa-8* gene cluster is carried by a pathogenicity island inserted into the tRNA<sup>Phe</sup> of human and bovine pathogenic *Escherichia coli* isolates. *Infect. Immun.* **69**:937–948.
  32. **Landraud, L., M. Gibert, M. R. Popoff, P. Boquet, and M. Gauthier.** 2003. Expression of *cnf1* by *Escherichia coli* J96 involves a large upstream DNA region including the *hlyCABD* operon, and is regulated by the RfaH protein. *Mol. Microbiol.* **47**:1653–1667.
  33. **Le Bouguéneq, C., M. Archambaud, and A. Labigne.** 1992. Rapid and specific detection of the *pap*, *afa*, and *sfa* adhesin-encoding operons in uropathogenic *Escherichia coli* strains by polymerase chain reaction. *J. Clin. Microbiol.* **30**:1189–1193.
  34. **Ochman, H., and R. K. Selander.** 1984. Standard reference strains of *Escherichia coli* from natural populations. *J. Bacteriol.* **157**:690–693.
  35. **Picard, B., P. Duriez, S. Gouriou, I. Matic, E. Denamur, and F. Taddei.** 2001. Mutator natural *Escherichia coli* isolates have an unusual virulence phenotype. *Infect. Immun.* **69**:9–14.
  36. **Picard, B., C. Journet-Mancy, N. Picard-Pasquier, and P. Goulet.** 1993. Genetic structures of the B2 and B1 *Escherichia coli* strains responsible for extra-intestinal infections. *J. Gen. Microbiol.* **139**:3079–3088.
  37. **Ritter, A., D. L. Gally, P. B. Olsen, U. Dobrindt, A. Friedrich, P. Klemm, and J. Hacker.** 1997. The PAI-associated *leuX* specific tRNA<sup>Leu</sup> affects type 1 fimbriation in pathogenic *Escherichia coli* by control of FimB recombinase expression. *Mol. Microbiol.* **25**:871–882.
  38. **Russo, T. A., and J. R. Johnson.** 2000. Proposal for a new inclusive designation for extraintestinal pathogenic isolates of *Escherichia coli*: ExPEC. *J. Infect. Dis.* **181**:1753–1754. Epub 15 May 2000.
  39. **Schubert, S., A. Rakin, and J. Heesemann.** 2004. The *Yersinia* high-pathogenicity island (HPI): evolutionary and functional aspects. *Int. J. Med. Microbiol.* **294**:83–94.
  40. **Swenson, D. L., N. O. Bukanov, D. E. Berg, and R. A. Welch.** 1996. Two pathogenicity islands in uropathogenic *Escherichia coli* J96: cosmid cloning and sample sequencing. *Infect. Immun.* **64**:3736–3743.
  41. **van Die, I., I. van Megen, W. Hoekstra, and H. Bergmans.** 1984. Molecular organisation of the genes involved in the production of F7(2) fimbriae, causing mannose-resistant haemagglutination, of a uropathogenic *Escherichia coli* 06:K2:H1:F7 strain. *Mol. Gen. Genet.* **194**:528–533.

---

Editor: J. B. Bliska