Commensal *Escherichia coli* isolates are phylogenetically distributed among geographically distinct human populations

Patrick Duriez,¹ Olivier Clermont,² Stéphane Bonacorsi,²
Edouard Bingen,² André Chaventré,³ Jacques Elion,¹
Bertrand Picard⁴ and Erick Denamur¹

Author for correspondence: Erick Denamur. Tel: +33 1 400 31916. Fax: +33 1 400 31903.
e-mail: denamur@infobiogen.fr

^{1,2} INSERM U458¹,
Laboratoire d'études de
génétique bactérienne
dans les infections de
l'enfant (EA3105)²,
Hôpital Robert Debré,
48 Boulevard Sérurier
75019 Paris, France

³ Département
d'Anthropologie
Génétique, Bordeaux II,
3 Place de la Victoire,
33000 Bordeaux, France

⁴ Laboratoire de
Microbiologie, Faculté
de médecine de Brest,
22 Avenue Camille
Desmoulins, B.P. 815,
29285 Brest Cedex, France

An intraspecies phylogenetic grouping of 168 human commensal *Escherichia coli* strains isolated from the stools of three geographically distinct human populations (France, Croatia, Mali) was generated by triplex PCR. The distributions of seven known extraintestinal virulence determinants (*ibeA*, *pap*, *sfa/foc*, *afa*, *hly*, *cnf1*, *aer*) were also determined by PCR. The data from the three populations were compiled, which showed that strains from phylogenetic groups A (40 %) and B1 (34 %) were the most common, followed by phylogenetic group D strains (15 %). Strains of the phylogenetic group B2 were rare (11 %). However, a significant specific distribution for strains of groups A, B1 and B2 within each population was observed, which may indicate the influence of (i) geographic/climatic conditions, (ii) dietary factors and/or the use of antibiotics or (iii) host genetic factors on the commensal flora. Virulence determinants were rarely detected, with only 25.6 % of the strains harbouring at least one of the virulence genes tested. The strains with virulence factors most frequently belonged to phylogenetic group B2. The commensal strains of phylogenetic groups A, B1 and D had fewer virulence determinants than pathogenic strains of the corresponding groups when these data were compared with those for previous collections of virulent extraintestinal infection strains studied using the same approach. However, the virulence patterns of commensal and pathogenic B2 phylogenetic group strains were the same. The data thus suggest that strains of the A, B1 and D phylogenetic groups predominate in the gut flora and that these strains must acquire virulence factors to become pathogenic. In contrast, commensal phylogenetic group B2 strains are rare but appear to be potentially virulent.

Keywords: *Escherichia coli*, phylogeny, commensal, virulence

INTRODUCTION

Escherichia coli is a normal inhabitant of the gut flora and is also the Gram-negative bacillus most frequently isolated in cases of human infection (Eisenstein & Zaleznick, 2000). Commensal enteric *E. coli* may therefore be the natural reservoir of pathogenic strains (Falkow, 1996). Intestinal or extraintestinal *E. coli* infections are caused by strains harbouring numerous virulence factors located on plasmids, bacteriophages, or the bacterial chromosome (Mühldorfer & Hacker, 1994). Several studies have shown that pathogenic *E. coli* strains may be derived from commensal strains by

the acquisition of chromosomal or extra-chromosomal virulence operons (Finlay & Falkow, 1997; Ochman *et al.*, 2000). 'Black hole' genomic deletions that enhance pathogenicity (Maurelli *et al.*, 1998) or random functional point mutations that are adaptive for pathogenic environments (Sokurenko *et al.*, 1998) may also represent additional strategies of genome plasticity by which a commensal strain may become virulent. Many of these observations were made possible by the fact that *E. coli* populations have a clonal structure (Selander & Levin, 1980). Four main phylogenetic groups, A, B1, B2 and D, were described by Herzer *et al.* (1990) using multilocus enzyme electrophoresis with the 72 strains of

the *E. coli* reference (ECOR) collection (Ochman & Selander, 1984). This finding was subsequently confirmed by Desjardins *et al.* (1995), by comparison of several genetic markers. The assignment of *E. coli* clones to one of these four groups is the basis of phylogenetic studies of the species. Phylogenetic analyses have indicated several origins for diverse natural populations of pathogenic *E. coli* and *Shigella* strains (Pupo *et al.*, 1997). Thus, *Shigella* clones are derived from *E. coli* outside the phylogenetic groups B2 and A (Rolland *et al.*, 1998; Pupo *et al.*, 2000), whereas verotoxin-producing *E. coli* O157:H7 clones could belong to phylogenetic group D (Clermont *et al.*, 2000). The clones responsible for human extraintestinal infections frequently belong to the anciently diverged B2 phylogenetic group (Lecointre *et al.*, 1998; Picard *et al.*, 1999).

The population structure of several pathogenic isolates has been extensively studied but little is known about the structure of commensal strain populations. It is therefore mandatory to determine the intraspecies phylogenetic relationships of *E. coli* isolates from the normal gut flora of healthy humans, to establish a background database for further studies on pathogenic strains. In this study, we used a recently described method facilitating the rapid and simple determination of *E. coli* phylogenetic group (Clermont *et al.*, 2000) in 168 non-epidemiologically related isolates from three geographically distinct human populations. This technique is based on a triplex PCR using a combination of two genes (*chuA* and *yjaA*) and an anonymous DNA fragment. The results obtained were strongly correlated with those obtained by multilocus enzyme electrophoresis and ribotyping methods (Clermont *et al.*, 2000). The distributions of several known extraintestinal virulence factors (*pap*, *afa*, *sfa/foc* adhesin-encoding operons, *hly* and *aer* operons and *ibeA* and *cnf1* genes) were also determined.

METHODS

Bacterial strains. One hundred and sixty-eight commensal strains of *E. coli* were obtained from faecal samples collected from 168 healthy unrelated human subjects between 1980 and 1990. The stools were cultured on Drigalski agar plates and one lactose-positive colony was randomly picked from each sample. *E. coli* were identified by an API 20E gallery (bioMérieux). Three sets of strains were obtained from three distinct geographic locations: 56 strains were from Paris, France; 57 strains were from Olib and Silba, two islands of Croatia; and 55 strains were from a Dogon population in Mali.

PCR grouping. The phylogenetic group to which the *E. coli* strains belonged was determined by a PCR-based method as described previously (Clermont *et al.*, 2000). Briefly, a two-step triplex PCR was performed directly on 3 µl bacterial lysate. The primer pairs used were *ChuA*.1 (5'-GACGAAC-CAACGGTCAGGAT-3')/*ChuA*.2 (5'-TGCCGCCAGTAC-CAAAGACA-3'), *YjaA*.1 (5'-TGAAGTGTCTCAGGAGACGCTG-3')/*YjaA*.2 (5'-ATGGAGAATGCGTTCCTCAAC-3') and *TspE4C2.1* (5'-GAGTAATGTCTGGGGCATTCA-3')/*TspE4C2.2* (5'-CGCGCCAACAAAGTATTACG-3'). The PCR steps were as follows: denaturation for 4 min at 94 °C, 30 cycles of 5 s at 94 °C and 10 s at 59 °C, and a final extension

step of 5 min at 72 °C. The data of the three amplifications resulted in assignment of the strains to phylogenetic groups as follows: *chuA* +, *yjaA* +, group B2; *chuA* +, *yjaA* -, group D; *chuA* -, *TspE4.C2* +, group B1; *chuA* -, *TspE4.C2* -, group A (Clermont *et al.*, 2000).

Detection of virulence factors. Virulence factors were detected by amplifying the corresponding gene (*ibeA*, *pap*, *afa*, *sfa/foc*, *hly*, *cnf1* and *aer*) from DNA by PCR as described previously (Picard *et al.*, 1999, 2001).

Statistical analysis. In addition to the data obtained experimentally in this work, the commensal strains were compared with two previously analysed collections of extraintestinal pathogenic strains (Bingen *et al.*, 1998; Picard *et al.*, 1999). The frequencies of various phylogenetic groups in the human populations were compared using the χ^2 test and the distributions of the number of virulence factors were compared using the Mann-Whitney *U*-test. All these tests were carried out using the Statview 4.5 software (Abacus Concept). A *P* value less than 0.05 was considered to show a significant difference.

RESULTS AND DISCUSSION

Distribution of the human commensal strains among the *E. coli* phylogenetic groups

When the strains recovered from all the populations were considered, A and B1 group strains were the most common (40 and 34 %, respectively), followed by D group strains (15 %). B2 group strains were rare (11 %) (Table 1). The rarity of B2 group strains confirms previous findings obtained with carboxylesterase B electrophoretic typing, which indicated that the carboxylesterase B₂ type [which corresponds to phylogenetic group B2 (Picard *et al.*, 1987)] accounts for 9 % of commensal human isolates and 1.6 % of animal isolates (Goullet & Picard, 1986a).

The proportion of D group strains did not differ between populations, whereas there was a significant (χ^2 test) increasing gradient of A group strains from Malian to French via Croatian populations (24–61 %) which correlates with a significant (χ^2 test) decreasing gradient of B1 group strains (58–12.5 %). Thus, a predominance of A strains in the French population and of B1 strains in the Malian population was observed. Our data also show that the proportion of B2 phylogenetic group strains differs between populations. These strains are found in Malian populations at a frequency similar to that in animals (2 %), whereas they account for almost one-fifth

Table 1. Distribution of the 168 human commensal *E. coli* strains among the four phylogenetic groups

Phylogenetic groups are as in Clermont *et al.* (2000). Data are given as numbers of strains (percentage in parentheses).

Population	A	B1	D	B2	Total
Mali	13 (24)	32 (58)	9 (16)	1 (2)	55
Croatia	20 (35)	18 (32)	8 (14)	11 (19)	57
France	34 (61)	7 (12.5)	9 (16)	6 (10.5)	56
Total	67 (40)	57 (34)	26 (15)	18 (11)	168

Table 2. Frequency of virulence genes among the commensal and pathogenic *E. coli* strains belonging to the four phylogenetic groups

Phylogenetic groups are as in Clermont *et al.* (2000). Data are given as numbers of strains (percentage in parentheses). The percentage is calculated as the number of strains with a given virulence determinant over the total number of strains belonging to each phylogenetic group $\times 100$ in the first four columns and over the total number of strains $\times 100$ in the last column. C, Commensal strains (this work); P, pathogenic strains from Bingen *et al.* (1998) and Picard *et al.* (1999).

Virulence gene	Origin of strain	Phylogenetic group				All groups
		A	B1	D	B2	
<i>aer</i>	C	8 (12)	3 (5.3)	6 (23)	10 (55.5)	27 (16)
	P	7 (64)	2 (67)	16 (84)	65 (76)	90 (76)
<i>pap</i>	C	2 (3)	3 (5.3)	5 (19.2)	9 (50)	19 (11.3)
	P	6 (55)	1 (33)	14 (74)	47 (55)	68 (58)
<i>afa</i>	C	1 (1.5)		1 (3.8)		2 (1.2)
	P	3 (27)		2 (11)		5 (4)
<i>sfa/foc</i>	C				6 (33.3)	6 (3.6)
	P				43 (51)	43 (36)
<i>hly</i>	C	1 (1.5)			6 (33.3)	7 (4)
	P				25 (29)	25 (21)
<i>cnf1</i>	C				6 (33.3)	6 (3.6)
	P				22 (26)	22 (19)
<i>ibeA</i>	C				3 (16.6)	3 (1.8)
	P				28 (33)	28 (24)

of the Croatian commensal *E. coli*. (Table 1). A previous study of *E. coli* commensal strains from animals and humans originating mainly from France and Africa, based on the electrophoretic polymorphism of esterases, showed a clear delineation between (i) French human commensal isolates and (ii) African human and animal isolates (Picard *et al.*, 1990). These differences in the distribution of the phylogenetic groups among the commensal strains of geographically different populations may be due to three main factors: (i) geographic/climatic conditions, (ii) dietary factors and/or the use of antibiotics, or (iii) host genetic factors; some *E. coli* strains may be primarily adapted to the gut conditions of certain populations. Analysis of the genetic structure of natural populations of *E. coli* in wild hosts on different continents suggest that strains are organized into an ecotypic structure in which adaptation to the host plays an important role in the population structure (Souza *et al.*, 1999).

Distribution of extraintestinal virulence genes among human commensal strains

Extraintestinal virulence genes were rarely detected among human commensal strains, as none of the seven virulence genes studied was detected in 74.4% of strains. Only 11.3, 1.2 and 3.6% of strains contained the *pap*, *afa* and *sfa/foc* adhesin-coding operons, respectively. The *aer* determinant was found in 16% of strains whereas the remaining virulence determinants were detected in less than 5% of strains (Table 2). These proportions differed according to the phylogenetic origin of the strains. The strains most frequently bearing virulence

factors were those belonging to phylogenetic group B2, in which two or three virulence genes were often detected per strain (Table 3). The 43 strains harbouring at least one virulence gene (25.6% of the 168 strains) included 28 harbouring one, 9 harbouring two, 2 harbouring three, 2 harbouring four and 2 harbouring five virulence genes (Table 3). The strains from Mali harboured fewer virulence factors than did the Croatian (*U*-test, $P = 0.005$) and French (*U*-test, $P = 0.0039$) isolates, as only one Malian strain (B2 group) possessed two virulence factors versus 23 strains in Croatia and 19 strains in France.

These data are consistent with most studies on the frequency of virulence factors in strains from healthy human stools (Minshew *et al.*, 1978; Hagberg *et al.*, 1981; Goullet & Picard, 1986b; Opal *et al.*, 1990; Siitonen, 1992; Hilali *et al.*, 2000). The frequent occurrence of commensal strains with virulence determinants reported by Mühldorfer *et al.* (1996) may be due to a local epidemiological bias, as no such prevalence was found by these workers in Peru (Hacker *et al.*, 1983).

Comparison of human commensal strain data with data obtained from two extraintestinal pathogen collections

The phylogenetic grouping of strains used for comparison were all obtained by the PCR triplex method. The 69 neonatal meningitis strains (Bingen *et al.*, 1998) and the 49 strains isolated from miscellaneous extraintestinal infections in adults (Picard *et al.*, 1999) were reanalysed by the PCR grouping method and the results

Table 3. The 43 *E. coli* commensal strains with at least one virulence gene

Strain*	Phylogenetic group	Virulence gene							No. of virulence genes detected
		<i>ibeA</i>	<i>cnf1</i>	<i>hly</i>	<i>sfa/foc</i>	<i>afa</i>	<i>pap</i>	<i>aer</i>	
c4689	A	—	—	—	—	—	—	+	1
c4705	A	—	—	—	—	—	+	—	1
c4720	A	—	—	—	—	+	—	—	1
c4735	A	—	—	—	—	—	—	+	1
c4752	A	—	—	—	—	—	—	+	1
vdg386	A	—	—	—	—	—	—	+	1
vdg388	A	—	—	—	—	—	—	+	1
vdg396	A	—	—	—	—	—	+	—	1
vdg422	A	—	—	—	—	—	—	+	1
vdg428	A	—	—	+	—	—	—	—	1
vdg433	A	—	—	—	—	—	—	+	1
vdg447	A	—	—	—	—	—	—	+	1
c4694	B1	—	—	—	—	—	+	—	1
c4722	B1	—	—	—	—	—	+	—	1
c4740	B1	—	—	—	—	—	—	+	1
vdg370	B1	—	—	—	—	—	—	+	1
c4712	D	—	—	—	—	—	—	+	1
c4714	D	—	—	—	—	—	—	+	1
vdg397	D	—	—	—	—	—	+	—	1
vdg405	D	—	—	—	—	—	+	—	1
vdg406	D	—	—	—	—	—	—	+	1
vdg457	D	—	—	—	—	—	+	—	1
c4695	B2	+	—	—	—	—	—	—	1
c4701	B2	+	—	—	—	—	—	—	1
c4719	B2	—	—	—	—	—	+	—	1
c4728	B2	—	—	—	—	—	—	+	1
c4739	B2	—	—	—	—	—	—	+	1
vdg427	B2	—	—	—	—	—	—	+	1
c4760	B1	—	—	—	—	—	+	+	2
c4692	D	—	—	—	—	—	+	+	2
c4731	D	—	—	—	—	+	—	+	2
vdg375	D	—	—	—	—	—	+	+	2
c4711	B2	—	—	—	—	—	+	+	2
c4713	B2	—	—	—	—	—	+	+	2
c4761	B2	—	—	—	—	—	+	+	2
m1391	B2	—	—	—	—	—	+	+	2
vdg425	B2	—	—	—	—	—	+	+	2
vdg430	B2	—	+	+	+	—	—	—	3
vdg448	B2	—	+	+	+	—	—	—	3
vdg383	B2	+	+	+	+	—	—	—	4
vdg395	B2	—	+	+	+	—	+	—	4
c4691	B2	—	+	+	+	—	+	+	5
c4737	B2	—	+	+	+	—	+	+	5

* The origin of the strains is as follows: c, Croatia; m, Mali; vdg, France.

obtained were strongly correlated with those of ribotyping (Clermont *et al.*, 2000 and data not shown). Virulence determinants were also detected using the same approach. The *cnf1* data for the pathogen collections are unpublished.

Of the 118 extraintestinal pathogenic strains, 110 (93 %) harboured at least one virulence gene. The human

commensal strains had fewer virulence factors than the extraintestinal pathogenic strains (*U*-test, $P < 0.0001$). This was also true for the frequencies of each individual virulence determinant, except for *afa*, which was under-represented (Table 2).

Pathogenic strains were distributed among phylogenetic groups as follows: A, 11 strains (9 %); B1, 3 strains

(2.5%); D, 19 strains (16%) and B2, 85 strains (72%). If virulence determinant frequencies were compared by phylogenetic group between the commensal and extra-intestinal pathogenic series, the commensal strains of phylogenetic groups A, B1 and D appeared to possess fewer virulence determinants (U -test, $P < 0.0001$, = 0.0017 and < 0.0001 , respectively). In contrast, no significant difference was observed between the phylogenetic B2 strains. This absence of difference was also found for each virulence gene (Table 3).

In addition, confirming that the strains of phylogenetic group B2 are more frequent among extraintestinal pathogenic strains (40–72%) than among commensal strains (9–11%) (Goullet & Picard, 1986b; Johnson *et al.*, 1991; Bingen *et al.*, 1998; Picard *et al.*, 1999; Hilali *et al.*, 2000; this work), our homogeneous dataset also increases our understanding of the emergence of virulence in human extraintestinal infections. The smaller number of virulence determinants in commensal strains of phylogenetic groups A, B1 and D than in the corresponding pathogenic strains is consistent with the hypothesis (Bingen *et al.*, 1998; Lecoindre *et al.*, 1998; Picard *et al.*, 1999) that these strains acquire virulence factors by horizontal transfer, enabling them to become virulent and to invade immunocompetent hosts (Picard & Goullet, 1988) from the host's intestinal reservoir. In contrast, the strains of phylogenetic group B2 in the commensal flora appear to be potentially virulent. This intrinsic virulence may explain why strains of phylogenetic group B2 are frequently isolated from patients with extraintestinal infections, despite their rarity in the gut flora.

Concluding remarks

This work on a relatively small number of commensal strains clearly demonstrates that the commensal flora has a complex population structure, which depends on ecotype. As commensal enteric isolates are believed to be the reservoirs of pathogenic strains, it will be interesting to determine whether the geographic differences between *E. coli* commensal populations are associated with differences in the distribution of diseases caused by different types of *E. coli*. Studies on a large number of commensal strains including precise geographic, socio-economic and medical data are needed to provide further insight into the emergence of virulent *E. coli* clones.

ACKNOWLEDGEMENTS

We thank Christine Amarin, Stéphanie Gouriou and Charlotte Tacheau for technical assistance. This work was supported in part by the Programme de Recherche Fondamentale en Microbiologie et Maladies Infectieuses et Parasitaires.

REFERENCES

- Bingen, E., Picard, B., Brahimi, N., Mathy, S., Desjardins, P., Elion, J. & Denamur, E. (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.
- Clermont, O., Bonacorsi, S. & Bingen, E. (2000). Rapid and simple determination of the *Escherichia coli* phylogenetic group. *Appl Environ Microbiol* **66**, 4555–4558.
- Desjardins, P., Picard, B., Kaltenböck, B., Elion, J. & Denamur, E. (1995). Sex in *Escherichia coli* does not disrupt the clonal structure of the population: evidence from random amplified polymorphic DNA and restriction-fragment-length polymorphism. *J Mol Evol* **40**, 440–448.
- Eisenstein, B. I. & Zaleznick, D. F. (2000). Enterobacteriaceae. In *Principles and Practice of Infectious Disease*, vol. 2, pp. 2294–2310. Edited by G. L. Mandell, J. E. Benet & R. Dolin. New York: Churchill Livingstone.
- Falkow, S. (1996). The evolution of pathogenicity in *Escherichia coli*, *Shigella*, and *Salmonella*. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*, 2nd edn, pp. 2723–2769. Edited by F. C. Neidhardt and others. Washington, DC: American Society for Microbiology.
- Finlay, B. B. & Falkow, S. (1997). Common themes in microbial pathogenicity revisited. *Microb Mol Biol Rev* **61**, 136–169.
- Goullet, P. & Picard, B. (1986a). Comparative esterase electrophoretic polymorphism of *Escherichia coli* isolates obtained from animal and human sources. *J Gen Microbiol* **132**, 1843–1851.
- Goullet, P. & Picard, B. (1986b). Highly pathogenic strains of *Escherichia coli* revealed by the distinct electrophoretic patterns of carboxylesterase B. *J Gen Microbiol* **132**, 1853–1858.
- Hacker, J., Schroter, G., Schrettenbrunner, A., Hugues, C. & Goebel, W. (1983). Hemolytic *Escherichia coli* strains in the human fecal flora as potential urinary pathogens. *Zentbl Bakteriol Mikrobiol Hyg* **254**, 370–378.
- Hagberg, L., Jodal, U., Korhonen, T. K., Lidin-Janson, G., Lindberg, U. & Svanborg Eden, C. (1981). Adhesion, hemagglutination, and virulence of *Escherichia coli* causing urinary tract infections. *Infect Immun* **31**, 564–570.
- Herzer, P. J., Inouye, S., Inouye, M. & Whittam, T. S. (1990). Phylogenetic distribution of branched RNA-linked multicopy single-stranded DNA among natural isolates of *Escherichia coli*. *J Bacteriol* **172**, 6175–6181.
- Hilali, F., Ruimy, R., Saulnier, P., Barnabé, C., Lebougué, C., Tibayrenc, M. & Andremon, A. (2000). Prevalence of virulence genes and clonality in *Escherichia coli* strains that cause bacteremia in cancer patients. *Infect Immun* **68**, 3983–3989.
- Johnson, J. R., Goullet, P., Picard, B., Moseley, S. L., Roberts, P. L. & Stamm, W. E. (1991). Association of carboxylesterase B electrophoretic pattern with presence and expression of urovirulence factor determinants and antimicrobial resistance among strains of *Escherichia coli* that cause urosepsis. *Infect Immun* **59**, 2311–2315.
- Lecoindre, G., Rachdi, L., Darlu, P. & Denamur, E. (1998). *Escherichia coli* molecular phylogeny using the incongruence length difference test. *Mol Biol Evol* **15**, 1685–1695.
- Maurelli, A. T., Fernandez, R. E., Bloch, C. A., Rode, C. K. & Fasano, A. (1998). 'Black holes' and bacterial pathogenicity: a large genomic deletion that enhances the virulence of *Shigella* spp. and enteroinvasive *Escherichia coli*. *Proc Natl Acad Sci U S A* **95**, 3943–3948.
- Minshew, B. H., Jorgensen, J., Counts, G. W. & Falkow, S. (1978). Association of hemolysin production, hemagglutination of human erythrocytes, and virulence for chicken embryos of extraintestinal *Escherichia coli* isolates. *Infect Immun* **20**, 50–54.
- Mühldorfer, I. & Hacker, J. (1994). Genetic aspects of *Escherichia coli* virulence. *Microb Pathog* **16**, 171–181.
- Mühldorfer, I., Blum, G., Donohue-Rolfe, A., Heier, H.,

- Ölschlger, T., Tschpe, H., Wallner, U. & Hacker, J. (1996). Characterization of *Escherichia coli* strains isolated from environmental water habitats and from stool samples of healthy volunteers. *Res Microbiol* **147**, 625–635.
- Ochman, H. & Selander, R. K. (1984). Standard reference strains of *Escherichia coli* from natural populations. *J Bacteriol* **157**, 690–693.
- Ochman, H., Lawrence, J. G. & Groisman, E. A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**, 299–304.
- Opal, S. M., Cross, A. S., Gemski, P. & Lythe, L. W. (1990). Aerobactin and alpha-hemolysin as virulence determinants in *Escherichia coli* isolated from human blood, urine, and stool. *J Infect Dis* **161**, 794–796.
- Picard, B. & Goulet, P. (1988). Correlation between electrophoretic types B₁ and B₂ of carboxylesterase B and host-dependent factors in *Escherichia coli* septicaemia. *Epidemiol Infect* **100**, 51–61.
- Picard, B., Picard-Pasquier, N. & Krishnamoorthy, R. (1987). Characterization of highly virulent *Escherichia coli* strains by ribosomal DNA restriction fragment length polymorphism. *FEMS Microbiol Lett* **82**, 183–188.
- Picard, B., Goulet, P., Sammartino, A. E. & Chaventr, A. (1990). Electrophoretic polymorphism of esterases in strains of *Escherichia coli* isolated from different host populations. In *Pluri-disciplinary Approach to Human Isolates*, pp. 233–242. Edited by the Institut National d’Etudes Dmographiques (INED), Paris.
- Picard, B., Sevali-Garca, J., Gouriou, S., Duriez, P., Brahimi, N., Bingen, E., Elion, J. & Denamur, E. (1999). The link between phylogeny and virulence in *Escherichia coli* extraintestinal infection. *Infect Immun* **67**, 546–553.
- Picard, B., Duriez, P., Gouriou, S., Matic, I., Denamur, E. & Taddei, F. (2001). Mutator natural isolates *Escherichia coli* isolates have an unusual virulence phenotype. *Infect Immun* **69**, 9–14.
- Pupo, G. M., Karaolis, R., Lan, R. & Reeves, P. R. (1997). Evolutionary relationship among pathogenic and non pathogenic *Escherichia coli* strains inferred from multilocus enzyme electrophoresis and *mdh* sequence studies. *Infect Immun* **65**, 2685–2692.
- Pupo, G. M., Lan, R. & Reeves, P. R. (2000). Multiple independent origins of *Shigella* clones of *Escherichia coli* and convergent evolution of many of their characteristics. *Proc Natl Acad Sci USA* **97**, 10567–10572.
- Rolland, K., Lambert-Zechovsky, N., Picard, B. & Denamur, E. (1998). *Shigella* and enteroinvasive *Escherichia coli* strains are derived from distinct ancestral strains of *E. coli*. *Microbiology* **144**, 2667–2672.
- Selander, R. K. & Levin, B. R. (1980). Genetic diversity and structure in *Escherichia coli* populations. *Science* **210**, 545–547.
- Siitonen, A. (1992). *Escherichia coli* in fecal flora of healthy adults: serotypes, P and type 1C fimbriae, non-P mannose-resistant adhesins, and hemolytic activity. *J Infect Dis* **166**, 1058–1065.
- Sokurenko, E. V., Chesnokova, V., Dykhuizen, D. E., Ofer, I., Wu, X. R., Krogfelt, K. A., Struve, C., Schembri, M. A. & Hasty, D. (1998). Pathogenic adaptation of *Escherichia coli* by natural variation of the FimH adhesin. *Proc Natl Acad Sci USA* **95**, 8922–8926.
- Souza, V., Rocha, M., Valera, A. & Eguiarte, L. E. (1999). Genetic structure of natural populations of *Escherichia coli* in wild hosts on different continents. *Appl Environ Microbiol* **65**, 3373–3385.

Received 25 October 2000; revised 1 February 2001; accepted 14 February 2001.