Evidence for a human-specific *Escherichia coli* clone

Olivier Clermont,[1] Mathilde Lescat,[1]
Claire L. O'Brien,[2] David M. Gordon,[2]
Olivier Tenaillon[1] and Erick Denamur[1]*

[1]*INSERM U722 and Université Paris 7 Denis Diderot, France.*
[2]*School of Botany and Zoology, The Australian National University, Canberra, ACT 0200, Australia.*

Summary

***Escherichia coli* is a widespread commensal of the vertebrate intestinal tract. Until recently, no strong association between a particular clone and a given host species has been found. However, members of the B2 subgroup VIII clone with an O81 serotype appear to be human host specific. To determine the degree of host specificity exhibited by this clone, a PCR-based assay was used to screen 723 faecal and clinical isolates from humans, and 904 faecal isolates from animals. This clone was not detected among the animal isolates, but was discovered in people living in Africa, Europe and South America. The clone is rarely isolated from people suffering from intestinal or extraintestinal disease and is avirulent in a mouse model of extraintestinal infection. Fine-scale epidemiological analysis suggests that this clone is competitively dominant relative to other members of the B2 phylogenetic group and that it has increased in frequency over the past 20 years. This clone appears to be a good candidate for use as a probiotic, and may be suitable as an indicator of human faecal contamination in microbial source tracking studies.**

Introduction

Escherichia coli is a widespread commensal of the lower intestinal tract of humans and other vertebrates that occasionally causes intestinal and extraintestinal diseases (Donnenberg, 2002). Understanding what makes *E. coli* an occasionally devastating pathogen requires a better knowledge of its ecology as a commensal. A first step in gaining this knowledge would be to determine if niche specialization occurs within the species. *E. coli* colonizes a wide variety of hosts with very diverse gut morphologies and digestive physiologies. *E. coli* also exhibits substan-

tial genetic structure, with most faecal isolates belonging to one of four genetic groups (A, B1, D and B2) (Herzer *et al.*, 1990; Wirth *et al.*, 2006). However, epidemiological studies have only demonstrated weak associations between host species and the prevalence of strains of the four phylogenetic groups (Gordon and Cowling, 2003; Escobar-Páramo *et al.*, 2006). Although the frequency of the four genetic groups varies to some degree with the diet or body mass of the host from which they are isolated, other factors such as climate, year of sampling, or the domestication status of the animals sampled (wild versus domesticated) also shape the genetic structure of *E. coli*. To date, no strong association between a particular clone and a given host species has been found, except perhaps for a *hly* B1 clone that appears to be restricted to animals (Escobar-Páramo *et al.*, 2006).

The great majority of *E. coli* strains belonging to genetic group B2 are highly virulent in a mouse model of extraintestinal infection (Johnson *et al.*, 2006). Recently, as part of a detailed determination of the genetic makeup of group B2 strains, we identified an unusual B2 strain, in that it is avirulent in the mouse model of extraintestinal infection (Le Gall *et al.*, 2007). This strain has an O81 serotype and is a member of the B2 clonal subgroup VIII as determined by multilocus sequence typing (MLST) (Le Gall *et al.*, 2007). Preliminary epidemiological evidence suggested that this strain might be human specific. Consequently, we undertook a large-scale epidemiological survey with the goal of determining the distribution of this strain in human and non-human vertebrates across four continents.

Results

Global distribution of the B2 VIII/O81 clone

A total of 1369 strains belonging to phylogenetic group B2, representing 438 faecal and 285 clinical isolates from humans, and 646 faecal isolates from wild, domesticated and zoo animals, were screened using the O81-specific primer. An additional 258 faecal isolates from non-domesticated animals and 15 O81 strains from septicemic chickens, all of unknown phylogenetic group, were also screened. All O81-positive strains were screened using the B2 clonal subgroup VIII-specific primers and, if unknown, their phylogenetic group membership was determined.

Despite surveying strains from over 60 animal species living on four continents, none of the animal isolates were

found to possess the O81 locus and to be members of B2 clonal subgroup VIII. By contrast, 64 members of the B2 VIII/O81 clone were detected among the human isolates. Two of the B2 VIII/O81 strains were clinical isolates; one strain (IAI48) was responsible for a urinary tract infection (Picard *et al.*, 1999) and the other (381A) was an entero-aggregative strain responsible for a case of diarrhoea (Escobar-Páramo *et al.*, 2004a). B2 VIII/O81 strains were isolated from humans living in Africa, the Americas and Europe, but not from people living in Australia.

Seven B2 strains from both non-human vertebrates and humans yielded an O81 polymerase chain reaction (PCR) product but were not members of the B2 clonal subgroup VIII, indicating that the O81 serotype can be found in other B2 clonal groups. Further, three O81-positive strains belonging to phylogenetic groups A and B1 were detected from mammals living in the Americas (1 group A and 2 group B1 strains), as well as 15 O81 B1 strains from chickens sampled in Spain. All of the O81 strains that were not members of the B2 VIII/O81 clone exhibited a polymorphic pattern of virulence determinants, that was, in every case, distinct from the B2 VIII/O81 clone (data not shown).

To assess degree of sequence type diversity among the B2 VIII/O81 strains, we choose six isolates from different hosts and geographical regions, as well as the two disease causing isolates (Table 1) for further characterization. We first performed MLST by sequencing six essential genes (5921 bp) in these strains. The sequences were all identical except for one strain (Ben4d) that has a synonymous polymorphism in the *putP* gene.

We also performed an *in silico* search for strains of the B2 VIII/O81 clone in the Max Planck Institute MLST database (http://web.mpiib-berlin.mpg.de/mlst/dbs/Ecoli). This database contained, at the time it was searched, data for 1404 disease and faecal isolates largely recovered from humans and domesticated animals. For this search, the allele sequence for each of the seven genes used in this MLST scheme (*adk*, *fumC*, *gyrB*, *icd*, *mdh*, *recA* and *purA*) was determined for the B2 VIII/O81 strain ED1a. Only one strain (M716, ST 452), an isolate from an Australian mammal, matched the profile of ED1a. However, this isolate does not yield an O81 PCR product, and M716 differs from ED1a at seven bases over the 727 nucleotides of *trpA* locus, including the allele-specific site of the *trpA*VIII.f primer, as well as in its virulence profile.

Virulence characteristics of B2 VIII/O81 strains

The eight B2 VIII/O81 strains were tested in an experimental mouse model of extraintestinal virulence. Whereas most B2 strains exhibit a high virulence in such a model (Johnson *et al.*, 2006), none of the tested strains

Table 1. Principal characteristics of strains representing the O81 B2 subgroup VIII clone.

Strain	Host origin	Condition of isolation	neuC	papC	papG	sfa/foc	hly	cnf1	iroN	ireA	aer	fyuA	irp2	iha	hra	sat	usp	ompT	ibeA	malX	traT
ED1a	France	Faeces[a]	–	–	–	–	–	–	–	+	+	+	+	+	–	–	+	+	–	+	–
VDG427	France	Faeces	–	–	–	–	–	–	–	+	+	–	–	+	–	–	+	+	–	+	–
Ben4d	Benin	Faeces	–	–	–	–	–	–	–	+	+	–	–	+	–	–	+	+	–	+	–
Ben27a	Benin	Faeces	–	–	–	–	–	–	–	+	+	+	+	+	–	–	+	+	–	+	–
colF12g	Colombia	Faeces	–	–	–	–	–	–	–	+	+	+	+	+	–	–	+	+	–	+	–
colH13c	Colombia	Faeces	–	–	–	–	–	–	–	+	+	–	–	+	–	–	+	+	–	+	–
381A	Central African Republic	Diarrhea (EAEC)	–	–	–	–	–	–	–	+	+	–	–	+	–	–	+	+	–	+	–
IAI48	France	Urinary tract infection	–	–	–	–	–	–	–	+	+	–	–	+	–	–	+	+	–	+	–

a. Commensal.
EAEC, enteroaggregative *E. coli.*

had any virulence. We then screened the isolates for the presence of 19 putative virulence determinants associated with extraintestinal disease. All the strains exhibited a similar virulence gene profile (Table 1). The iron metabolism genes *ireA*, *aer* and *iha* were present in all strains. However, the genes *fyuA* and *irp2*, which are markers for the high-pathogenicity island, were absent from some strains. None of the genes classically associated with intraintestinal virulence was present in the ED1a reference strain.

Fine-scale analysis of the epidemiologic data

In a sample of faecal isolates collected post 2000 from 152 people living in Colombia, Benin, and France, and considering one randomly selected isolate per individual, the B2 VIII/O81 clone represented 4% of all isolates (3–7% according to the population) and 18% of the B2 strains recovered (11–28% according to the population). By contrast, in a sample of faecal isolates collected in the 1980s from 193 people living in the USA, Mali, Croatia, France and Sweden, the B2 VIII/O81 clone represented only 0.4% of all isolates tested and 4% of the B2 strains. This result suggests that the VIII/O81 clone has increased its representation in human faeces 10-fold in the past 20 years.

   To determine if the B2 VIII/O81 clone exists at a high frequency within a host we analysed the data set cited above, but used all 10 isolates taken from each of the 152 subjects. O81 was detected in 12 of the 152 hosts and, in 42% of these people, the O81 clone represented at least 80% of the *E. coli* isolates recovered from that host. By contrast a non-VIII/O81 B2 strain was detected in 58 hosts, but in these cases the B2 strain represented > 80% of the *E. coli* in that host only 17% of the time. Put differently, 77% of the O81 strains detected were found in hosts in which they were the majority clone (> 80%), whereas only 37% of non-O81 B2 strains were found in a situation where they represented the majority (> 80%) of the *E. coli* in a host. This outcome suggests that the B2 VIII/O81 clone is able to achieve numerical dominance in a human more often than other group B2 strains. Furthermore, in one subject, all 30 isolates taken from faeces emitted at three different times (10 isolates per faeces) over a 1-year period as well as 10 isolates from stools obtained 6 years after the first sample were the B2 VIII/O81 clone. This result demonstrates that this clone can persist in the human gut for extend periods.

Survival in water

A human-specific strain of *E. coli* could be a useful marker in microbial source-tracking studies that attempt to deter-

mine if human faecal contamination of a water body is occurring. Consequently, we tested the survival of ED1a and compared it with the survival of *E. coli* K12 (Bogosian *et al.*, 1996). In non-sterile river water at 20°C the decline in viable cell counts of both strains are identical, and by day 9 neither strain can be detected (Fig. 1A). This result is in accord with previously reported outcomes for K12 (Bogosian *et al.*, 1996). In sterile river water at 20°C, viable cell counts of both strains declined by only 1–2 log in 35 days, with ED1a exhibiting poorer survival than K12 (Fig. 1B). These findings are consistent with earlier suggestions that the decline of *E. coli* populations in water is due to predation by protozoa, phages and exposure to heat-labile toxins.

Discussion

Host-specific *E. coli* pathogens are not unknown, for example it has long been thought that *Shigella* strains are primate specific (Donnenberg, 2002) and there are rabbit-specific enteropathogenic *E. coli* (Leyton *et al.*, 2007). However, this is the first report of what is apparently a human host-specific commensal strain of *E. coli*, although it is not known if this strain can colonize other primates. This clone belongs to the B2 phylogenetic group and, although B2 strains are often responsible for extraintestinal infection (Picard *et al.*, 1999), strains of this clone appear to rarely cause disease. Only two pathogenic strains belonging to this clone have been isolated so far, one strain causing a urinary tract infection and an entero-aggregative strain. The B2 VIII/O81 clone is absent from the Acthman *E. coli* MLST database and over half of the more than 1400 isolates in this database are listed as pathogens (Wirth *et al.*, 2006). Based on its virulence determinant pattern and/or serotype, this clone is also absent from European and American collections of B2 strains responsible for newborn meningitis (*n* = 99, Bonacorsi *et al.*, 2003) and urosepsis (*n* = 108, Johnson and Stell, 2000; Bingen-Bidois *et al.*, 2002). Finally, none of the strains of this clone is virulent in the mouse model of extraintestinal infection.

   B2 VIII/O81 clone appears to be highly successful as it has been found on three continents and, in some populations, it can represent up to 7% of the isolates taken from human faeces. The reason for the absence of this clone in humans living in Australia is unknown. Geographic isolation is unlikely to be the answer, as over 60% of the MLST sequence types isolated from Australian humans have been detected in other parts of the world (D. Gordon unpubl. data). All isolates of the B2 VIII clone with an O81 serotype are also very homogeneous in that they possess a very similar suit of genes, while in other B2 strains these genes have a highly polymorphic distribution (this work and Le Gall *et al.*, 2007). These observations

A

B
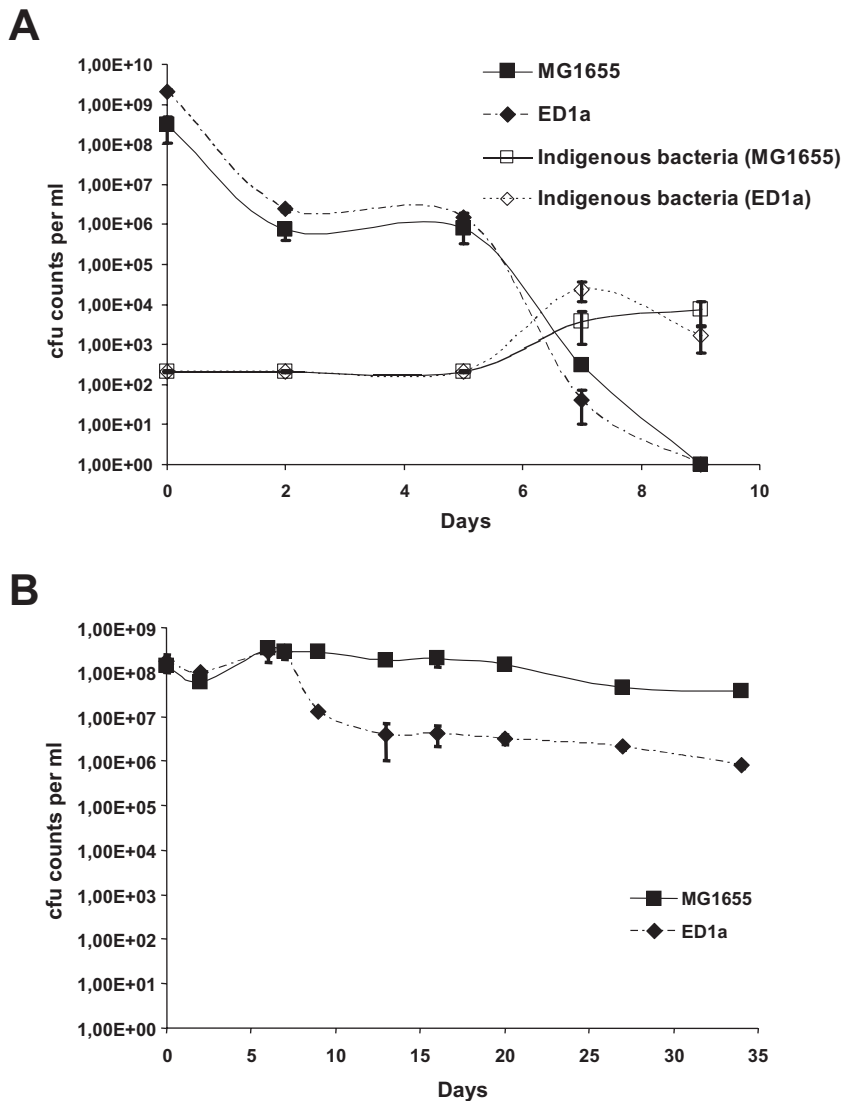
Fig. 1. Decline of K12-MG1655 and ED1a *E. coli* cells in (A) non-sterile and (B) sterile water of the Saint Martin canal in Paris at 20°C. Indigenous bacteria counts are indicated in the non-sterile experiment.

are consistent with the idea that the clone has undergone a recent expansion.

The existence of a human-specific *E. coli* clone has potential applied applications. The B2 VIII/O81 clone could be used as a probiotic, as it appears to be avirulent and good colonizer of the human gut. The avirulent nature of this strain means that it might be particularly appropriate for use in highly vulnerable patients, such as those in intensive care, the immunocomprised, or premature infants. By comparison, the currently used probiotic strains, Nissle O6 strain (Mutaflor®) isolated in 1917 from the faeces of a healthy soldier (Sun *et al.*, 2005) and the A0 34/86 O83 strain isolated from porcine faeces (Hejnova *et al.*, 2005), are highly virulent in our mouse model of extraintestinal virulence. Nissle O6 killed all 10 inoculated mice and A0 34/86 O83 killed nine of the 10 mice inoculated.

Microbial source tracking could be another application of the O81/B2 subgroup VIII clone. Coliforms and *E. coli*, in particular, have been used for many years, to assess both water quality and to determine the source of any faecal pollution that might occur in a water body (Scott *et al.*, 2002). However, the use of *E. coli* in these efforts has been hampered by the apparent lack of any extensive host specificity in *E. coli* (Gordon, 2001) and by the ability of many *E. coli* strains to undergo significant cell division in the external environment (Barnes and Gordon, 2004; Power *et al.*, 2005). The usefulness of the microbial indicators as a tool in risk assessment can be significantly enhanced by the identification of a human-specific clone. The fact that not all individuals carry the O81/B2 subgroup VIII clone is not restrictive, as human faecal pollution is generally due to release of sewage from wastewater treatment plants or septic tanks, and therefore represents

faecal contamination not by individuals but by populations. The inability of the B2 subgroupVIII/O81 clone to undergo significant cell division in water makes it a good candidate marker of human faecal contamination of water. Furthermore, the PCR assay that we have developed can be used to identify the clone directly from water or soil samples.

We have identified an *E. coli* clone that is an excellent candidate for further molecular investigation aimed at understanding (i) what defines a commensal *E. coli* strain and (ii) what attributes of the strain are responsible for its host specificity. Consequently, we have initiated a project whose aim is to sequence the genome of strain ED1a (http://www.genoscope.cns.fr/externe/English/Projets/ Projet_MW/organisme_MW.html) in conjunction with an in-depth metabolic characterization of this strain.

Experimental procedures

Bacterial strains

Several previously published collections were studied. A collection of 293 isolates representing all of the B2 strains among 1520 isolates taken from the faeces of 152 humans (10 strains per individual) (Escobar-Páramo *et al.*, 2004b). These strains originated from Benin (68 isolates/24 subjects), Colombia (81 isolates/18 subjects) and France (144 isolates/25 subjects) and were obtained post 2000. Twenty-five human faecal B2 isolates from 25 subjects originating from Europe (France, Sweden, Croatia), Africa (Mali) and America (USA) and sampled in the 1980s (Ochman and Selander, 1984; Duriez *et al.*, 2001) were also studied. A collection of 378 faecal isolates (310 isolates from 87 mammals, 68 isolates from 23 birds) representing all of the B2 strains among 1930 isolates from the faeces of 405 animals (one to 10 strains per individual). These strains were sampled between 1980 and 2000 in France, Africa (Ethiopia, Cameroon, Gabon) and the Americas (French Caribbean Islands, Venezuela, USA) from domesticated and wild (Escobar-Páramo *et al.*, 2006), as well as zoo (Ochman and Selander, 1984) animals. A collection of 55 pathogenic B2 isolates from humans (one strain per patient) encompassing enteropathogenic, enteroaggregative, diffusely adhering and extraintestinal pathovars (Ochman and Selander, 1984; Picard *et al.*, 1999; Escobar-Páramo *et al.*, 2004a). These strains originated from France, Sweden, Brazil, USA and Central African Republic, and were collected between 1980 and 2000. The collection of strains isolated from humans living in the Canberra region of Australia was acquired by isolating a single colony from the faeces or an extraintestinal body site of each person sampled (Gordon *et al.*, 2005; Gordon and O'Brien, 2006). A total of 120 B2 faecal isolates and 230 B2 strains taken from extraintestinal body sites were examined for this study. The collection of strains from native wild Australian vertebrates was acquired by selecting a single isolate from each faecal sample and a single faecal sample was taken from each host individual (Gordon and Cowling, 2003). The numbers of Australian B2 strains examined were: mammals ($n = 234$), birds ($n = 30$) and reptiles ($n = 4$). A collection described by Souza and colleagues (1999) and con-

sisting of 258 *E. coli* faecal isolates from animals was analysed. For the Souza collection a single isolate per host was taken from faecal material of the following host types: mammals ($n = 197$), birds ($n = 51$) and reptiles ($n = 10$) living primarily in Central and South America. Lastly, 15 O81 *E. coli* strains isolated from the hearts of 15 chickens suffering from septicaemia in Spain during the 1990s (Blanco *et al.*, 1997) were also studied.

PCR screening of O81and B2 subgroup VIII strains

The serotype of the B2 clonal subgroup VIII strain ED1a was determined via traditional serology (Dr Chobi Deb Roy, *E. coli* reference center, Pennsylvania State University, USA). To design an O81-specific primer pair, the approach described by Clermont and colleagues (2007) was used. First, the 5′ extremity of the *rfb* cluster from ED1a was sequenced and using this sequence data an O81-specific primer was designed (rfbO81bisr: 5′-GAGCAGTATATATTACTGGTG-3′). This primer, together with the gndbis.f primer (5′-ATACC GACGACGCCGATCTG-3′), yields a 383 bp fragment.

Primers to confirm that a strain was a member of the B2 clonal subgroup VIII were designed for the *trpA* gene. The forward primer (trpAVIII.f: 5′-GCGCAAAGAAGGCGCATTCA-3′) is specific to B2 subgroup VIII strains and together with a non-specific *trpA* reverse primer (trpA2.r: 5′-GCAACGCGG CCTGGCGGAAG-3′) produces a 394 bp fragment. The universal forward primer targeting the 3′ portion of *trpB* (trpBA.f: 5′-CGGCGATAAAGACATCTTCA-3′) which, together with the trpA2.r primer, yields a 489 bp product from all *E. coli* strains was included in the reaction as a positive control.

Polymerase chain reactions were carried out in a 20 µl volume and contained 2 µl of 10× buffer (supplied with *Taq* polymerase), 20 pmol of each primer, 2 µM each dNTP, 1 U of *Taq* polymerase (Ozyme, St Quentin en Yvelines, France), and 3 µl of bacterial lysate. Polymerase chain reactions were performed with an Eppendorf Mastercycler with MicroAm tubes using the following amplification conditions: 4 min at 94°C, 30 cycles of 5 s at 94°C and 10 s at 59°C, with a final extension step of 5 min at 72°C. Polymerase chain reaction products were loaded on 2% agarose gels containing SYBR® Safe DNA gel stain (Invitrogen, Cergy Pontoise, France) and visualized and photographed under UV light.

MLST analysis

The phylogenetic relationships among the strains were inferred using nucleotide sequence data from six essential genes [*trpA*, *trpB*, *pabB*, *putP*, *icd*, and *polB* (Gerdes *et al.*, 2003)] which are thought to experience little recombination and produce a strong phylogenetic signal (Escobar-Páramo *et al.*, 2004c).

Virulence gene detection

The following 19 genes, usually considered as extraintestinal 'virulence determinants' and involved in capsule synthesis (*neuC*), adhesion (*papC*, *papG*, *sfa/foc*), toxin production (*hly*, *cnf1*), iron metabolism (*iron*, *ireA*, *aer*, *aer*, *fyuA*, *irp2*,

iha) and other traits (*hra*, *sat*, *usp*, *ompT*, *ibeA*, *malX*, *traT*), were detected by using established simplex or multiplex PCR assays (Johnson *et al.*, 2006). An *in silico* search of the complete genome of B2 subgroup VIII strain ED1a (V. Barbe, C. Médigue, E. Denamur unpubl. data) was undertaken to determine if this strain harboured genes also found in the main pathovars of diarrhoeagenic *E. coli*, i.e. diffusely adhering *E. coli* (*afaBC/daaC*), enteroaggregative *E. coli* (AA probe), enterohemorragic *E. coli* and enteropathogenic *E. coli* (*eae*), enteropathogenic *E. coli* (*bfpA*), enterotoxigenic *E. coli* (STa and LT-1), enterohemorragic *E. coli* (*stx1*, *stx2*, and *ehxA*) and enteroinvasive *E. coli* (*ipaB*) (Escobar-Páramo *et al.*, 2004a).

Survival of *E. coli* strains in water

The river water used was obtained from the Saint Martin canal in Paris at a site about 1 km upstream from the Seine. For the sterile water studies, the water was autoclaved for 45 min. About 100 ml aliquots of water were placed into 250 ml Erlenmeyer flasks capped with paper covers. After inoculation, the flasks were incubated at 20°C under a 12:12 light : dark cycle. Indigenous bacteria were present at a level of about 2×10^2 cfu ml^{-1} of river water. Fresh cultures of the strains K12-MG1655 and ED1a that have been grown for 14 h in 10 ml of Luria–Bertani (LB) medium were washed with sterile 0.9% saline, centrifuged, re-suspended in 1 ml and added to the water. Samples for counting were removed directly from the water, diluted and plated on Drigalsky and LB (non-sterile water) or LB (sterile water) media. In the non-sterile water experiment, the differentiation between indigenous and inoculated bacteria was based on colony morphology and O81 PCR assay. All experiments were carried out in duplicate.

Mouse lethality assay

A mouse model of systemic infection was used to assess the intrinsic virulence of the O81 strains (Picard *et al.*, 1999). For each strain, 10 outbred female Swiss OF1 mice (3–4 weeks old, 14–16 g) were challenged subcutaneously in the abdomen with a standardized bacterial inoculum (10^9 cfu ml^{-1} of log-phase bacteria in 0.2 ml Ringer solution). Mortality was assessed over 7 days post challenge. The urosepsis strains CFT073 was used as a positive control and the faecal-derived strain K12-MG1655 as a negative control. Data were available for three strains (ED1a, Ben4d, IAI48) (Le Gall *et al.*, 2007). In this model system, lethality is a rather clear-cut parameter and strains were usually classified as non-killer (strains killing < 2 mice out of 10) or killer (strains killing > 8 mice) (Johnson *et al.*, 2006).

Acknowledgements

References

Barnes, B., and Gordon, D. (2004) Coliform dynamics and the implications for source tracking. *Environ Microbiol* **6:** 501–509.

Bingen-Bidois, M., Clermont, O., Bonacorsi, S., Terki, M., Brahimi, N., Loukil, C., *et al.* (2002) Phylogenetic analysis and prevalence of urosepsis strains of *Escherichia coli* bearing pathogenicity island-like domains. *Infect Immun* **70:** 3216–3236.

Blanco, J.E., Blanco, M., Mora, A., and Blanco, J. (1997) Production of toxins (enterotoxins, verotoxins, and necro-toxins) and colicins by *Escherichia coli* strains isolated from septicemic and healthy chickens: relationship with in vivo pathogenicity. *J Clin Microbiol* **35:** 2953–2957.

Bogosian, G., Sammons, L.E., Morris, P.J.L., O'Neil, J.P., Heitkamp, M.A., and Weber, D.B. (1996) Death of the *Escherichia coli* K-12 strain W3110 in soil and water. *Appl Environ Microbiol* **62:** 4114–4120.

Bonacorsi, S., Clermont, O., Houdouin, V., Corevant, C., Brahimi, N., Marecat, A., *et al.* (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.

Clermont, O., Johnson, J.R., Menard, M., and Denamur, E. (2007) Determination of *Escherichia coli* O types by allele-specific polymerase chain reaction: application to the O types involved in human septicemia. *Diagn Microbiol Infect Dis* **57:** 129–156.

Donnenberg, M. (2002) *Escherichia coli* virulence mechanisms of versatile pathogen. Elsevier Science Ed ed, San Diego, CA, USA.

Duriez, P., Clermont, O., Bonacorsi, S., Bingen, E., Chaventre, A., Elion, J., *et al.* (2001) Commensal *Escherichia coli* isolates are phylogenetically distributed among geographically distinct human populations. *Microbiology* **147:** 1671–1676.

Escobar-Páramo, P., Clermont, O., Blanc-Potard, A., Bui, H., Le Bouguenec, C., and Denamur, E. (2004a) A specific genetic background is required for acquisition and expression of virulence factors in *Escherichia coli*. *Mol Biol Evol* **21:** 1085–1094.

Escobar-Páramo, P., Grenet, K., Le Menach, A., Rode, L., Salgado, E., Amorin, C., *et al.* (2004b) Large-scale population structure of human commensal *Escherichia coli* isolates. *Appl Environ Microbiol* **70:** 5698–5700.

Escobar-Páramo, P., Sabbagh, A., Darlu, P., Pradillon, O., Vaury, C., Denamur, E., and Lecointre, G. (2004c) Decreasing the effects of horizontal gene transfer on bacterial phylogeny: the *Escherichia coli* case study. *Mol Phylogenet Evol* **30:** 243–250.

Escobar-Páramo, P., Le Menac'h, A., Le Gall, T., Amorin, C., Gouriou, S., Picard, B., *et al.* (2006) Identification of forces shaping the commensal *Escherichia coli* genetic structure by comparing animal and human isolates. *Environ Microbiol* **8:** 1975–1984.

Gerdes, S.Y., Scholle, M.D., Campbell, J.W., Balazsi, G., Ravasz, E., Daugherty, M.D., *et al.* (2003) Experimental determination and system level analysis of essential genes in *Escherichia coli* MG1655. *J Bacteriol* **185:** 5673–5684.

Gordon, D.M. (2001) Geographical structure and host specificity in bacteria and the implications for tracing the source of coliform contamination. *Microbiology* **147:** 1079–1085.

Gordon, D.M., and Cowling, A. (2003) The distribution and genetic structure of *Escherichia coli* in Australian verte-

brates: host and geographic effects. *Microbiology* **149:** 3575–3586.

Gordon, D.M., and O'Brien, C.L. (2006) Bacteriocin diversity and the frequency of multiple bacteriocin production in *Escherichia coli. Microbiology* **152:** 3239–3244.

Gordon, D.M., Stern, S.E., and Collignon, P.J. (2005) Influence of the age and sex of human hosts on the distribution of *Escherichia coli* ECOR groups and virulence traits. *Microbiology* **151:** 15–23.

Hejnova, J., Dobrindt, U., Nemcova, R., Rusniok, C., Bomba, A., Frangeul, L., *et al.* (2005) Characterization of the flexible genome complement of the commensal *Escherichia coli* strain A0, 34/86 (O83: K24: H31). *Microbiology* **151:** 385–398.

Herzer, P.J., Inouye, S., Inouye, M., and 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.

Johnson, J.R., and Stell, A.L. (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.

Johnson, J.R., Clermont, O., Menard, M., Kuskowski, M.A., Picard, B., and Denamur, E. (2006) Experimental mouse lethality of *Escherichia coli* isolates, in relation to accessory traits, phylogenetic group, and ecological source. *J Infect Dis* **194:** 1141–1150.

Le Gall, T.O., Clermont, S., Gouriou, B., Picard, X., Nassif, E., Denamur and Tenaillon, O. (2007) Extraintestinal virulence is a coincidental by-product of commensalisms in B2 phylogenetic group *Escherichia coli* strains. *Mol Biol Evol* **24:** 2373–2384. August 19. Epub ahead of print.

Leyton, D.L., Adams, L.M., Kelly, M., Sloan, J., Tauschek, M., Robins-Browne, R.M., and Hartland, E.L. (2007) Contribution of a novel gene *rpeA* encoding a putative autotransporter adhesin to intestinal colonization by rabbit enteropathogenic *Escherichia coli. Infect Immun* **75:** 4664–4669.

Ochman, H., and Selander, R.K. (1984) Standard reference strains of *Escherichia coli* from natural populations. *J Bacteriol* **157:** 690–693.

Picard, B., Garcia, J., Gouriou, S., Duriez, P., Brahimi, N., Bingen, E., *et al.* (1999) The link between phylogeny and virulence in *Escherichia coli* extraintestinal infection. *Infect Immun* **67:** 546–553.

Power, M.L., Littlefield-Wyer, J., Gordon, D.M., Veal, D.A., and Slade, M.B. (2005) Phenotypic and genotypic characterisation of encapsulated *Escherichia coli* isolated from blooms in two Australian lakes. *Environ Microbiol* **7:** 631–640.

Scott, T.M., Rose, J.B., Jenkins, T.M., Farrah, S.R., and Lukasik, J. (2002) Microbial source tracking: current methodology and future directions. *Appl Environ Microbiol* **68:** 5796–5803.

Souza, V., Rocha, M., Valera, A., and Eguiarte, L.E. (1999) Genetic structure of natural populations of *Escherichia coli* in wild hosts on different continents. *Appl Environ Microbiol* **65:** 3373–3385.

Sun, J., Gunzer, F., Westendorf, A.M., Buer, J., Scharfe, M., Jarek, M., *et al.* (2005) Genomic peculiarity of coding sequences and metabolic potential of probiotic *Escherichia coli* strain Nissle 1917 inferred from raw genome data. *J Biotechnol* **117:** 147–161.

Wirth, T., Falush, D., Lan, R., Colles, F., Mensa, P., Wieler, L.H., *et al.* (2006) Sex and virulence in *Escherichia coli*: an evolutionary perspective. *Mol Microbiol* **60:** 1136–1151.