Phylogenetic, virulence and antibiotic resistance characteristics of commensal strain populations of Escherichia coli from community subjects in the Paris area in 2010 and evolution over 30 years.

Méril Massot, Anne-Sophie Daubié, Olivier Clermont, Françoise Jauréguy, Camille Couffignal, Ghizlane Dahbi, Azucena Mora, Jorge Blanco, Catherine Branger, France Mentré, et al.

Revision MIC-D-15-00380

Phylogenetic, virulence and antibiotic resistance characteristics of commensal strain populations of *Escherichia coli*from community subjects in the Paris area in 2010 and evolution over 30years

Running title: Commensal human *E. coli* in Paris

Authors and affiliations: Méril Massot[1,2,*], Anne-Sophie Daubié[1,3,4,*], Olivier Clermont[1,2,*], Françoise Jauréguy[1,3,4],Camille Couffignal[1,2], Ghizlane Dahbi[5],Azucena Mora[5],Jorge Blanco[5], Catherine Branger[1,2], France Mentré[1,2], Alain Eddi[6], Bertrand Picard[1,3,4], Erick Denamur[1,2], and the COLIVILLE Group[†]

[*] These authors have equally contributed to the work

[1] INSERM, IAME, UMR1137, Paris, France

[2] Univ Paris Diderot, IAME, UMR1137, Sorbonne Paris Cité, Paris, France

[3] APHP, Hôpitaux Universitaires Paris Seine Saint-Denis, Site Avicenne, Bobigny, France

[4] Univ Paris Nord, IAME, UMR1137, Sorbonne Paris Cité, Bobigny, France

[5]Laboratorio de Referencia de *E. coli* (LREC), Departamento de Microbioloxía e Parasitoloxía, Facultade de Veterinaria, Universidade de Santiago de Compostela (USC), Lugo, Spain

[6] Département de Médecine Générale, Univ Paris Diderot, Sorbonne Paris Cité, Faculté de Médecine, Paris, France

[†]The COLIVILLE group is composed of the following medical practitioners who were involved in the recruitment of subjects for the study: Monique Allouche, Jean-Pierre Aubert, Isabelle Aubin, Ghislaine Audran, Dan Baruch, Philippe Birembaux, Max Budowski, Emilie Chemla, Alain Eddi, Marc Frarier, Eric Galam, Julien Gelly, Serge Joly, Jean-François Millet, Michel Nougairede, Nadja Pillon, Guy Septavaux, Catherine Szwebel, Philippe Vellard, Raymond Wakim, Xavier Watelet and Philippe Zerr.

Corresponding author: Erick Denamur, IAME, UMR1137, INSERM, Université Paris Diderot, Site Xavier Bichat, 16, rue Henri Huchard, 75018 Paris, France. Tel: 33 1 57 27 77 39; email: erick.denamur@inserm.fr

Key words:*Escherichia coli*, phylogroup, microbiota, commensal

Content category: Environmental and Evolutionary Microbiology

Word counts:294words in the abstract,4019in the main text

SUMMARY

It is important to study commensal populations of *Escherichia coli*because they appear to be the reservoir of both extra-intestinal pathogenic*E. coli* and antibiotic resistant strains of *E. coli*.We studied 279dominant faecal strains of*E. coli*from 243adults living in thecommunity in the Paris area in 2010. The phylogenetic group and sub-group [sequence type complex (STc)]of the isolates and the presence of 20 virulence genes were determined by PCR assays.The O-types and the resistance to 18 antibiotics were assessed phenotypically. The B2 group was the most frequently recovered (34.0%), followed by the A group (28.7%),andother groups were rarer. The most prevalent B2 subgroups were II (STc73), IV (STc141), IX (STc95) and I (STc131) with22.1%, 21.1%, 16.8% and 13.7%,respectively,of the B2 group strains. Virulence factors (VFs) were more common in B2group than other strains.One or more resistance was found in 125 strains (44.8%of the collection) but only six (2.2% of the collection) weremultiresistant; no extended-spectrum beta-lactamase-producing strain was isolated. The C phylogroup and clonal group A strains were the most resistant. No trade-off between virulence and resistance was evidenced.Wecompared these strains to collections of strains gathered in the same conditions 30 and 10 years ago. There has been a parallel and linked increase in the frequency of B2 group strains (from 9.4% in 1980 to 22.7% in 2000 and 34.0% in 2010)and of VFs. Antibiotic resistancealso increased, from 22.6% of strains resistant to at least one antibiotic in 1980 to 31.8% in 2000 and 44.8% in 2010; resistance to streptomycin, however,remained stable.Commensal human *E. coli* populations have clearly evolved substantially over time, presumably reflecting changes in human practices, and particularlyincreasing antibiotic use.

INTRODUCTION

Escherichia coli is a gut commensal of vertebrates, including humans(Berg, 1996). It can nevertheless cause a broad range of diseases from various diarrheal diseases to extra-intestinal diseases, and particularlyurinary tract infections and bacteraemia(Russo & Johnson, 2003, Kaper *et al.*, 2004). The gut is considered to be the reservoir of strainscausing these extra-intestinal pathologies, and both epidemiological and experimental studies have led to the notion that virulence could be a by-product of commensalism(Nowrouzian *et al.*, 2005, Le Gall *et al.*, 2007, Diard *et al.*, 2010).

Molecular typing,based onmultilocus sequence typing (MLST), various PCR based approaches and, more recently, whole genome sequencing, hasprovideda good understanding of the phylogenetic framework of the *E. coli* species(Clermont *et al.*, 2015). The species *E. colisensu strict*ois largely clonal(Desjardins *et al.*, 1995), with sevenmain phylogenetic groups (A, B1, B2, C, D, E, and F) each composed of severalclones/clonal complexes or sub-groups(Clermont *et al.*, 2013). *Escherichia* clades have been described: they are phenotypically indistinguishable from *E. coli* but the nucleotide sequences are highly divergent, with clade I strains being the most closely related to *E. coli sensu stricto*(Walk *et al.*, 2009).Membership of these phylogenetic entities is of particular interest, as there is a relationship between the genetic background of astrain and itsvirulence factors(VFs) (Escobar-Paramo *et al.*, 2004). For example, some sequence types (STs) within the B2 phylogroup, asST137, includestrainswith numerous extra-intestinal genes and high virulence potential(Messika *et al.*, 2012).

Recently, there have been some major shifts in the epidemiology of *E. coli*. Some clonal complexes,such as ST complex (STc)69, STc73, STc95 and STc131,have become more prevalent in extra-intestinal diseases(Bidet *et al.*, 2007, Bert *et al.*, 2010, Mahjoub-Messai *et al.*, 2011, Bengtsson *et al.*, 2012, Gibreel *et al.*, 2012, Alhashash *et al.*, 2013, Clermont *et al.*, 2014). Some of theseclonal complexes have members that spread resistance to various antibiotics such as the cotrimoxazole [clonal group A (CGA) or STc69](Manges *et al.*, 2001)and third generation cephalosporins [ST131 producing extended-spectrum beta-lactamase (ESBL)](Nicolas-Chanoine *et al.*, 2014).The gut has been found to be at the hub of this resistance (Carlet, 2012) and a majordriver of the spread of antibiotic resistance in the community (Woerther *et al.*, 2013).

In viewof these major changes to*E. coli* epidemiology, it would be valuable to have recent and reliable data on human commensal strains. Unfortunately, such data are scarce and most relevant studies are based on specific populations, like hospitalised patients or travellers, and many are focused on resistant strains.We report a study of a collection of 279*E. coli* strains gathered from adult subjects living in the community in the Paris area(hereafter referred to as the "COLIVILLE" collection). We compared these strains to collections of strains gathered in the same conditions 30(Duriez *et al.*, 2001) and 10(Escobar-Paramo *et al.*, 2004)years ago by our group.

MATERIALS AND METHODS

Subjects of the COLIVILLE collection

Subjects were recruited by general practitioners of the Department of General Practice of the University Paris Diderot from the region of Ile-de-France (Paris, France, and its suburban area) from May 2009 to December 2011. All participants lived in the community and volunteered to self-collect a faecal swab sample. The inclusion criteria were: age of 18 years or more, no history of gastrointestinal disease, no symptoms of immunosuppression, no antibiotic therapy in the previous month and no hospitalisation in the 3 months preceding inclusion. Written informed consent was obtained from each participant, and the study was approved by the ethics evaluation committee of Institut National de la Santé et de la Recherche Médicale (INSERM) (CCTIRS no. 09.243, CNIL no. 909277, and CQI no. 01-014).

Isolationand characterisation of the COLIVILLE strains

Faecal samples were self-collected by the subjects. Immediatelyafter stool emission, a swab was dippedinto the faeces,put in Amies transport medium (Medical Wire & Equipment Corsham, Wiltshire England) and sent by mail to the Avicenne hospital laboratory (Bobigny, France)(Smati *et al.*, 2013). The swabs were plated onto Drigalski agar plates upon arrival (Bio-Rad, Life Science, Marnes-la-Coquette, France). After 24 hours of incubation at 37°C, the plates were inspected for the phenotypic aspect of the lactose-positive colonies. In most cases, only one type of colony was recovered. However, in some cases, two co-dominant types of colony were retrieved. One colony, or if appropriatetwophenotypically distinct colonies, were randomly picked, allowing the recovery of the dominant *E. coli* clones of

thecommensal faecal microbiota from each subject(Smati *et al.*, 2013). These isolates were identified as *E. coli* using API 20E (bioMérieux, Marcy l'Etoile, France), and stored in glycerol stock solution at -80°C.

***E. coli* genotyping**

DNA was extracted from colonies of each strain using the Wizardgenomic DNA purification kit (Promega, France) following the manufacturer instructions. The *E. coli*phylogroup (A, B1, B2, C, D, E, F) and*Escherichia* clade of each isolate were determined by the PCR quadruplex (Clermont *et al.*, 2013)and *aes* and *chuA* allele-specific amplification (Clermont *et al.*, 2011) methods, respectively. PCR-based methodswere used to identify strains of CGA clonal complex (STc69) within the D phylogroup (Johnson *et al.*, 2004)and to classify the B2 phylogroup strains into the 10 main subgroups (I to X)(Clermont *et al.*, 2014, Clermont *et al.*, 2015). Strains were screened for the presence of 20VFs representative of the main classes of *E. coli* extraintestinal virulence determinants known, including adhesins (*iha*, *papC*, *hra*,*sfa/foc*, *papGII* and *papGIII*alleles of *papG*, and *ibeA*), iron capture systems (*irp2*, *fyuA*, *iucC*,*iroN*and *ireA*), protectins (*ompT*, *traT* and *neuC*, chromosomal) and toxins (*usp*,*sat*,*clbQ*, *hlyC*, and *cnf1*) (Johnson *et al.*, 2008, Clermont *et al.*, 2011). For each isolate, a virulence score, defined as the number of the 20 VFs tested that were present in that strain, was calculated [adapted from (Lefort *et al.*, 2011)].

O-typing

Classical determination of O antigens was carried out according to Guinée & Jansen (1981), with all O (O1 to O181) antiseraavailable. Subtyping of O2, O6, O25 and O45 types was performed by PCR as previously described (Clermont *et al.*, 2011).

Antibiotic resistance phenotypes

The strains were tested for their antibiotic susceptibilities using adisk diffusion method according to the recommendations of the Antimicrobial Committee of the French Society for Microbiology(http://www.sfm-microbiologie.org). The following antimicrobial agents were tested: amoxicillin, amoxicillin-clavulanic acid, cefoxitin, cefotaxime, ertapenem, imipenem, streptomycin, gentamicin, tobramycin, netilmicin, amikacin, tetracycline, cotrimoxazole (sulfamethoxazole-trimethoprim), chloramphenicol, nalidixic acid, ofloxacin, fosfomycin, and nitrofurantoin. A strain was considered to be multiresistant ifit was resistant to,at least,penicillins, cotrimoxazole and quinolones (Lefort *et al.*, 2011). For each isolate, a

resistance score was computed. The resistance score was defined as the number of antibiotic classes to which thestrain was resistant, amongthe ten tested (penicillins, cephalosporins, carbapenems, aminoglycosides, tetracyclines, sulfonamides, amphenicols, quinolones, phosphonic acids, furans).

Characteristics of other collections of commensal strains

Two previously published collections of human commensal strains isolatedfrom the Paris area were studied for comparison. The criteria of selection of the subjects andprotocols of *E. coli* isolation were similar to those of the COLIVILLE collection, except for the number of isolatesstudied for eachindividual. The VDG collection was from subjects sampledin 1980 and included53 strains from 53 subjects (Duriez *et al.*, 2001)(onestrain per individual); the AEM collection wasobtainedin 2000 and included44 strains from 27 subjects(Escobar-Paramo *et al.*, 2004) (1 to 4 strainsper individual). The phylogroup/subgroup membership, the presence of 10 VFs representative of the four main classes (*sfa/foc*, *papC*, *iroN*, *iucC*, *fyuA*, *neuC*, *traT*, *hlyC*, *cnf1* and*usp*), and the susceptibility to seven antibiotics (amoxicillin, amoxicillin-clavulanic acid, streptomycin, chloramphenicol, cotrimoxazole,tetracycline andnalidixic acid)of these strains were determinedas above. For all the strains of the three collections, we calculated a 'reduced' virulence score, defined as the number of the 10 tested VFs present in the strain.

Statistical analyses

Study of the COLIVILLE strain determinants

Several bacterial determinants were compared between phylogroups using Fisher exact tests and analyses of variance. Fisher exact tests were usedto test associations between phylogroup and O-type and between phylogroup and VFs. Associations were tested for the most numerous O-types (10 or more strains), and all the studied VFs. Associations between the most frequent antibiotic resistances (at least five resistant strains) were similarly tested.

To evaluate the strength of association betweeneach VF and B2 group strains,Fisher's exact test was used to compare the distribution of each VF between B2 group andall the other groups combined. As multiple testswere performed, the p values were adjusted using the Benjamini and Hochberg method(Benjamini & Hochberg, 1995).Virulence and resistance scores were compared between phylogroups by analyses of variance; if significant, Tuckey's test was used for pairwise comparisons between phylogroups. Virulence scores were compared between B2 subgroups in a similar way. The Pearson correlation test was used to

assess the relationship between virulence and resistance scores.The resistance scores of the clonal group A strains and the other phylogroup Dstrainswere compared using a Wilcoxon-Mann-Whitney test.

Comparison with other commensal collections

The proportion of phylogenetic groupsand ofresistant strains were compared betweenthe three Parisian collections (VDG, AEM andCOLIVILLE) by Fisher exact tests. The 'reduced' virulence scores of the three collectionswere compared using a linear regression model with the date of sampling and the phylogenetic group as predictors.

All statistical analyses were performed using R software (R version 3.0.2). Data are shown as means with standarddeviations (SD) for continuous variables and number and percentagefor categorical variables. All tests were two-sided with a 5% type I error.

RESULTS

COLIVILLE subject population and strains

Two hundred and forty four subjects fulfilling our inclusion criteria were enrolled; there were 111 menand 133women. The mean age was 57 years (11.4): 57 years for the men (12.5) and 56 years for the women(10.3). The dominant *E. coli* clone was isolated from the stools of these participants by plating on selective Drigalski medium. *E. coli*colonies were recoveredfrom all subjects. For208 subjects, colonies were phenotypically homogeneousand one was randomly isolated for study. For 36 subjects, there were two phenotypically distinct types of colonies, and one representative of each type was isolated and studied.Thus, 280 strains were isolated and phenotypically identified as *E. coli* using API 20E; 279 from 243 subjects were identified as *E. coli sensu stricto*by the quadruplex PCR assay (Clermont *et al.*, 2013) and one strain from a subject exhibiting phenotypically homogeneous colonies was identified as *Escherichia* clade IV (Clermont *et al.*, 2011)(Table S1). This last subject was not included in the subsequent analyses.

Phylogenetic group/subgroup, O-type and VF distribution

The PCR quadruplex method was used to classify all isolates among the seven main phylogroups.Group B2 was the most frequently recovered (34.0%), followed by group A (28.7%) and the group B1 (12.9%); members ofgroups F (9.7%), D (9.0%), C (3.2%) and E (2.5%) were less numerous(Table 1).Many of the group B2 strains were infour sub-groups: sub-group II (STc73) (22.1%of the B2 group), sub-group IV (STc141) (21.1%), sub-group IX

(STc95) (16.8%) and sub-group I (STc131) (13.7%) (Table 1). Four sub-group Istrains belonged to ST131. Only 7.3% of the B2 strains were not assigned to a sub-group byour PCR-based assay (Clermont *et al.*, 2014). Forty-eight % of the D strains belonged to CGA (STc69).

The standard O-typing assay failed to assign 22 strains to an O-type. Among the 257 typeable strains, 75 O-types were detected, 28 represented by single strains and 47 by at least two strains.There were at least 10 strains in each of five O-types: O6a, O1, O2b, O9 and O18.There was an association between O-type and phylogenetic group:most of the O6a, O1 and O2b strains belonged to the B2, F and B2 groups, respectively (p value= $2x10^{-9}$),and mostnon-typable strains belonged to the A group.Three ofthe ST131 strains belonged to the ST131-O25b clone andone belonged to the ST131-O16 clone. Most strains of the B2 sub-groups II, IV, VI, VII and VIII were O6a-, O2b-, O4-, O75-, and O81-types, respectively, whereas the B2 subgroup I and IX strains were more heterogeneous (I: O25b, O9; IX: O1, O18, O2a). Inthe phylogroupD, CGA strainswere O17 and O77 types, and these typeswere not detected amongnon-CGA strains (Table S1).

The 279 strains were screened for 20 VF genes.VF genes involved in iron acquisition and protectins were more frequent than those for adhesins and toxins (Table 2). Three VF genes were foundinmore than half of the strains: *irp2* (60.9%) and *fuyA* (60.2%) both belonging to the high pathogenicity island (HPI) and involved in iron acquisition (Schubert *et al.*, 2009), and the protectin gene*ompT* (58.4%) (Table 2).The carriage of each of the 20 tested virulence factorsdifferedsignificantly between phylogroups (global adjusted p values≤ 0.005, Table 2). Seventeen VF genes were significantly more prevalent among phylogroup B2 strains than among other strains (adjusted p values≤0.002), whereas the three other (*iha*, *traT* and *sat*) were not(Table 2). Three VFs (*papGIII*, *clbQ* and *cnf1*) were found exclusively in the B2 strains.

The virulence score for the strains wasfrom 0 to 17, with a mean of 5.9(4.5) (Table 1). Virulence scorediffered significantly between the phylogroups (p value=10^{-16}). The virulence score of the B2 strains was significantly higher than those of the A, B1, C, D, E and F strains (p value= 10^{-13}, 10^{-13}, 10^{-6}, 10^{-13},10^{-8} and 10^{-7}, respectively). D strains had a significantly higher virulence score than A and B1 strains (p value= 0.004 and 0.01, respectively) and F strains had a significantly higher virulence score than A strains (p value= 10^{-10}). The virulence score wassignificantly different between subgroups of the B2 group (p value=10^{-5}): sub-groups II, IV and IX hadsignificantly higher virulence scores, with means of12.5 (2.3), 10.7

(3.0) and 11.2 (1.3), respectively, than the sub-group I withmean 7.2 (1.9) (p value= 10^{-5}, 0.004 and 0.001, respectively).

Antibiotic resistance phenotype

Resistance to one or more of the 18 tested antibioticswas found in 125 (44.8%) of the 279 strains (Table 1). Only 6 strains(2.2%) were multiresistant, *i.e.* resistant to, at least, amoxicillin, cotrimoxazole and nalidixic acid.Overall, 29.7% of the strains were resistant to amoxicillin and 16.1% to amoxicillin-clavulanic acid (Table 3). The only strain resistant tocefotaxime did not produce an ESBL but had a plasmid-encoded CMY-2 cephalosporinase (data not shown). No strain was resistant to carbapenems. Among the aminoglycosides, 23.3% of the strains were resistant to streptomycin and only 0.7%, 0.7% and 0.4% of the strains were resistant to gentamicin, tobramycin and netilmicin, respectively(Table 3). Resistance to tetracycline was found in 21.9% of the strains, to cotrimoxazole in 15.8% and to chloramphenicolin 7.2%;6.1% of strains were resistant to nalidixic acid and 3.2% to ofloxacin.Significant associations were found between penicillin resistance and resistance to the following antibiotics: chloramphenicol, cotrimoxazole, streptomycin, and tetracycline (adjusted p values<0.05).Associations between resistances to chloramphenicol, cotrimoxazole, streptomycin, and tetracycline were also significant (adjusted p values<10^{-3}). Resistance to nalidixic acid was significantly associated withresistance to amoxicillin-clavulanic acid, streptomycin, cotrimoxazole and chloramphenicol (p value= 0.049, 0.048, 0.049 and 0.007, respectively).

Resistance scores were between0and6, with a mean of 1.1(1.5)(Table 1). There was a significant difference of the resistance score between phylogroups (p value= 0.001). The resistance score of the C strains wassignificantly higher than those of the A, B1, B2, E and F strains (p value= 0.002, 0.02, 0.002, 0.03 and 0.03, respectively). The mean resistance score of the group D strainswas 1.7(2.0); the mean resistance score of CGA strains was2.7(1.7) and was significantly higher than the mean score of other group D strains, which was 0.9(1.8) (p value=0.009).

For the 279 strains as a whole, there was no significantcorrelation (Pearson coefficient= 0.06, p value=0.3; adjusted R^2 of linear regression= 0.0006, p value= 0.3) between the virulence and the resistance scores.

Comparison with other collectionsof commensal*E.coli*

We compared the COLIVILLE collection to collections obtained in 1980 (VDG) and 2000 (AEM) in the same area. The most apparent observationwas the increase in the proportion of group B2 strains from 9.4% in 1980 to 22.7% in 2000 and 34.0% in 2010, associated with a parallel decrease in the proportion of phylogroup A strains (52.8%, 27.3% and 28.7%, respectively, p value= 0.0002) (Fig. 1a).No CGA strain was evidenced in the VDG and AEM collections.The 'reduced'virulence score has increased significantly over time, froma mean of 1.5 (2.0) in 1980, to 2.3 (2.2) in 2000 and 3.2 (2.6) in 2010 (Fig. 1b). Our model indicatedthat this increase was due to the change in the phylogenetic distributionbetween 1980 and 2010 (effects of date of sampling were not significant), and in particular the increase in the proportion of the group B2 strains and the decrease in that of the A group strains. A more thorough analysis of the few B2 strains available from the VDG and AEM collections showed that classical extra-intestinal pathogenic *E. coli*(ExPEC) B2 subgroup strains (subgroup II/O6a and O22, subgroup III/O6a and subgroup IX/O1) with numerous VFs were already present (Table S2).

We also studied the evolution of antibiotic resistance in these populations (Fig. 1c). The proportion of strains resistant to at least one antibiotic increased from 22.6% in 1980 to 31.8% in 2000 and 44.8% in 2010(p value=0.005). We observed theemergence of the resistance to amoxicillin and amoxicillin-clavulanic acid to attain29.7% and 16.1%, respectively,by 2010.A similar trend was observed for cotrimoxazole and tetracycline whereas the resistance to streptomycin remained generally stable.

DISCUSSION

A major issue when isolating commensal strains ofan opportunistic pathogenic species, such as *E. coli*, is the quality of the epidemiological data availableforthe human host. We therefore applied very strict inclusion criteria for the subjects, excluding individuals with any antibiotic therapy in the previous month and any hospitalisation in the 3 months preceding inclusion; the subjects were all epidemiologically unrelated and were recruited in the same geographic area and overa short period of time. To our knowledge, this is the largest seriesof "true" human commensal strains to be thoroughly studied, with data including population genetics, virulence and antibiotic resistance data. This collection will be useful to the scientific community for subsequent comparison purposes (Table S1).

The high frequency of the phylogroup B2 strains in 2010 (Table 1) isconsistent with recent descriptions of human commensal strains in industrialised countries (Bailey *et al.*, 2010, Bailey *et al.*, 2010, Tenaillon *et al.*, 2010). It contrasts with data obtained in developing countries where phylogroup A and B1 strains predominate (Unno *et al.*, 2009, Bailey *et al.*, 2010, Li *et al.*, 2010, Tenaillon *et al.*, 2010, Lescat *et al.*, 2013, Tapader *et al.*, 2014).Group A strains are the bestadapted to a wide range of human hosts, whatever their lifestyle, diet and hygiene status.Ourstudy is one of the first to use the quadruplex PCR assay (Clermont *et al.*, 2013) allowing the identificationof the minor phylogenetic groups C, E and F.Few group C and E strains were found (3.2%and 2.5%, respectively) whereas almost 10% of strainsweregroup F, such that they were more numerous than group D strains (9.0%). Half ofthe group D strains were CGA strains,a clonal group includingExPEC(Manges *et al.*, 2001).Only one *Escherichia* clade was found, confirming the rarity of other*Escherichia* clades in human (Clermont *et al.*, 2011). We did not find any significant correlation between the phylogroup distribution and the age and orsex of the subjects (data not shown),contraryto observations in Australia(Gordon *et al.*, 2005).Wecompared our 2010 collection to previous collections from the Paris area.The frequency of isolation of phylogroup B2 strains has increased with a parallel decrease ofthat of phylogroup A strains (Fig. 1a); this may be a consequence of modifications in the lifestyle of the Parisian population including changing food processing and hygiene procedures. Conversely, the prevalence of the minor groups (C, E and F) has remained stable over time and maybe less sensitive to these changes.The presence of VFs is linked to phylogenetic group membership, a well-known phenomenon (Escobar-Paramo *et al.*, 2004), and we observed an increase of the number of VFs with time, paralleling the changes in phylogroup distribution (Fig. 1b).

Group B2 has various sub-groups or clonal complexes, three of them (sub-group I=STc131; sub-group II=STc73; sub-group IX=STc95) containing most ExPEC strains (Bengtsson *et al.*, 2012, Gibreel *et al.*, 2012, Alhashash *et al.*, 2013, Clermont *et al.*, 2014).We found numerous strains of sub-groups II and IX among the commensal B2 strains, with the classical ExPEC O-types (*i.e.* sub-group II/O6a and sub-group IX/O1-O18-O2a) whereas fewsub-group I strains, also frequently recovered among commensal strains, were the classical ST131/O25b-O16 clones (Johnson *et al.*, 2014).We also found many B2 sub-group IV (STc141) strains exhibiting the O2b type. It is therefore possible that two types of B2 strains in the human microbiotacan be distinguished: the classical ExPEC clones that have both the property of being highly virulent and the property of being good gut colonizers, the extra-intestinal

virulence being a by-product of commensalism (Nowrouzian *et al.*, 2005, Le Gall *et al.*, 2007, Diard *et al.*, 2010);and thecommensal-specific strains, as has been reported for the O81 B2 sub-group VIII (STc452) strains (Clermont *et al.*, 2008).

Overall, the strains we isolated exhibited a low level of antibiotic resistance (Table 1) relative to French clinical strains isolated over the same period (http://www.onerba.org). For example, resistances to amoxicillin, cefotaxime, gentamicin, cotrimoxazole and fluoroquinolones were 51.8%, 7.6%, 5.6%, 23.9% and 14.4%, respectively, for clinical strains in 2010, but only 29.7%, 0.4%, 0.7%, 15.8%, 6.1%, respectively, in the COLIVILLE collection.This low prevalenceof resistance in our collection argues indirectly for the "true" commensal character of the strains.Beside the well-known CGA(Manges *et al.*, 2001), our work identifies, for the first time, phylogroup C as apotential reservoir of resistance among commensal strains(Table 1). A strain of this phylogroupproducing an ESBL was responsible foran extensive outbreak of digestive colonisation in a neonatal ward with one case of meningitis (Moissenet *et al.*, 2010).

Although antibiotic resistance in our 2010 collection was less prevalent than amongclinical strains (see above), it was higher than in the earlier similar collections (Fig. 1c). This is consistent witha recent study in a Parisiancheck-up centre where a 10-fold increase of ESBL-producing *E. coli*strains was observed between 2006 and 2011(Nicolas-Chanoine *et al.*, 2013). These effects are probably due to an increase in the exposure of thegeneral population to antibiotic pressure (Woerther *et al.*, 2013).Contrasting with this trend of increasingresistance, the prevalence of resistance to streptomycin appears to be stable through time in the Parisianpopulation. A similar level of resistance to streptomycin (23%) was reported for*E. coli* strains isolated in England in 1991 (Chiew *et al.*, 1998). Streptomycin therapy was frequent in the 1950s and 1960s but is not now widely used, and it largely restricted to tuberculosis. Thenon-disappearance of resistance to an (almost) unused antibiotic may be because the genes conferring this resistance are borne by genetic structures,including transposons and integrons, that possess other antibiotic resistance genes (Chiew *et al.*, 1998).Indeed,we found an association in our collectionbetween resistance to streptomycin and resistance to cotrimoxazole and chloramphenicol, all these resistances being mediated by co-localised integron-borne genes(Partridge *et al.*, 2009).

In conclusion, our data indicates that commensal human *E. coli* populationsevolve substantially over time and that precise and repeated characterisations are requiredto documentthe emergence of virulent and/or resistant strains in the community.

ACKNOWLEDGMENTS

This work was partially supported by the grant CN2012/303(Consellería de Cultura, Educación e Ordenación Universitaria, Xunta de Galicia and The European Regional Development Fund, ERDF).

ABBREVIATIONS

CGA: Clonal Group A, ExPEC: Extraintestinal Pathogenic *Escherichia coli*, STc: Sequence Type complex, ST: Sequence Type, VF: Virulence Factor, MLST: Multilocus Sequence Typing, ESBL: Extended Spectrum Beta-Lactamase, SD: Standard Deviation, HPI: High Pathogenicity Island.

REFERENCES

Alhashash F, Weston V, Diggle M & McNally A (2013) Multidrug-resistant *Escherichia coli* bacteremia. *Emerg Infect Dis***19**: 1699-1701.

Bailey JK, Pinyon JL, Anantham S & Hall RM (2010) Commensal *Escherichia coli* of healthy humans: a reservoir for antibiotic-resistance determinants. *J Med Microbiol***59**: 1331-1339.

Bailey JK, Pinyon JL, Anantham S & Hall RM (2010) Distribution of human commensal *Escherichia coli* phylogenetic groups. *J Clin Microbiol***48**: 3455-3456.

Bengtsson S, Naseer U, Sundsfjord A, Kahlmeter G & Sundqvist M (2012) Sequence types and plasmid carriage of uropathogenic *Escherichia coli* devoid of phenotypically detectable resistance. *J Antimicrob Chemother***67**: 69-73.

Benjamini Y & Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol***57**: 289-300.

Berg RD (1996) The indigenous gastrointestinal microflora. *Trends Microbiol***4**: 430-435.

Bert F, Johnson JR, Ouattara B, Leflon-Guibout V, Johnston B, Marcon E, Valla D, Moreau R & Nicolas-Chanoine MH (2010) Genetic diversity and virulence profiles of *Escherichia*

coli isolates causing spontaneous bacterial peritonitis and bacteremia in patients with cirrhosis. *J Clin Microbiol***48**: 2709-2714.

Bidet P, Mahjoub-Messai F, Blanco J, Dehem M, Aujard Y, Bingen E & Bonacorsi S (2007) Combined multilocus sequence typing and O serogrouping distinguishes *Escherichia coli* subtypes associated with infant urosepsis and/or meningitis. *J Infect Dis***196**: 297-303.

Carlet J (2012) The gut is the epicentre of antibiotic resistance. *Antimicrob Resist Infect Control***1**: 39.

Chiew YF, Yeo SF, Hall LM & Livermore DM (1998) Can susceptibility to an antimicrobial be restored by halting its use? The case of streptomycin versus *Enterobacteriaceae*. *J Antimicrob Chemother***41**: 247-251.

Clermont O, Gordon D & Denamur E (2015) Guide to the various phylogenetic classification schemes for *Escherichia coli* and the correspondence among schemes. *Microbiology***161**: 980-988.

Clermont O, Christenson JK, Denamur E & Gordon DM (2013) The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. *Environ Microbiol Rep***8**: 58-65.

Clermont O, Gordon DM, Brisse S, Walk ST & Denamur E (2011) Characterization of the cryptic *Escherichia* lineages: rapid identification and prevalence. *Environ Microbiol***13**: 2468-2477.

Clermont O, Christenson JK, Daubie AS, Gordon DM & Denamur E (2014) Development of an allele-specific PCR for *Escherichia coli* B2 sub-typing, a rapid and easy to perform substitute of multilocus sequence typing. *J Microbiol Methods***101**: 24-27.

Clermont O, Lescat M, O'Brien CL, Gordon DM, Tenaillon O & Denamur E (2008) Evidence for a human-specific *Escherichia coli* clone. *Environ Microbiol***10**: 1000-1006.

Clermont O, Olier M, Hoede C, Diancourt L, Brisse S, Keroudean M, Glodt J, Picard B, Oswald E & Denamur E (2011) Animal and human pathogenic *Escherichia coli* strains share common genetic backgrounds. *Infect Genet Evol***11**: 654-662.

Desjardins P, Picard B, Kaltenbock 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***41**: 440-448.

Diard M, Garry L, Selva M, Mosser T, Denamur E & Matic I (2010) Pathogenicity-associated islands in extraintestinal pathogenic *Escherichia coli* are fitness elements involved in intestinal colonization. *J Bacteriol***192**: 4885-4893.

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

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

Escobar-Paramo P, Grenet K, Le Menac'h A, *et al.* (2004) Large-scale population structure of human commensal *Escherichia coli* isolates. *Appl Environ Microbiol***70**: 5698-5700.

Gibreel TM, Dodgson AR, Cheesbrough J, Fox AJ, Bolton FJ & Upton M (2012) Population structure, virulence potential and antibiotic susceptibility of uropathogenic *Escherichia coli* from Northwest England. *J Antimicrob Chemother***67**: 346-356.

Gordon DM, Stern SE & Collignon PJ (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.

Guinée PAM & Jansen WH (1981) *Escherichia coli* associated with neonatal diarrhoea in piglets and calves. *Laboratory diagnosis in neonatal calf and pig diarrhoea, current topics in veterinary and animal science*,(Leeww PM & Guinée PAM, eds.), p. 126-162. Martinus-Nijhoff, The Hague.

Johnson JR, Owens K, Manges AR & Riley LW (2004) Rapid and specific detection of *Escherichia coli* clonal group A by gene-specific PCR. *J Clin Microbiol***42**: 2618-2622.

Johnson JR, Johnston B, Kuskowski MA, Nougayrede JP & Oswald E (2008) Molecular epidemiology and phylogenetic distribution of the *Escherichia coli pks* genomic island. *J Clin Microbiol***46**: 3906-3911.

Johnson JR, Clermont O, Johnston B, Clabots C, Tchesnokova V, Sokurenko E, Junka AF, Maczynska B & Denamur E (2014) Rapid and specific detection, molecular epidemiology, and experimental virulence of the O16 subgroup within *Escherichia coli* sequence type 131. *J Clin Microbiol***52**: 1358-1365.

Kaper JB, Nataro JP & Mobley HL (2004) Pathogenic *Escherichia coli*. *Nat Rev Microbiol***2**: 123-140.

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

503 Lefort A, Panhard X, Clermont O, Woerther PL, Branger C, Mentre F, Fantin B, Wolff M &
504 Denamur E (2011) Host factors and portal of entry outweigh bacterial determinants to predict
505 the severity of *Escherichia coli* bacteremia. *J Clin Microbiol***49**: 777-783.
506 Lescat M, Clermont O, Woerther PL, *et al.* (2013) Commensal *Escherichia coli* strains in
507 Guiana reveal a high genetic diversity with host-dependant population structure. *Environ*
508 *Microbiol Rep***5**: 49-57.
509 Li B, Sun JY, Han LZ, Huang XH, Fu Q & Ni YX (2010) Phylogenetic groups and
510 pathogenicity island markers in fecal *Escherichia coli* isolates from asymptomatic humans in
511 China. *Appl Environ Microbiol***76**: 6698-6700.
512 Mahjoub-Messai F, Bidet P, Caro V, Diancourt L, Biran V, Aujard Y, Bingen E & Bonacorsi
513 S (2011) *Escherichia coli* isolates causing bacteremia via gut translocation and urinary tract
514 infection in young infants exhibit different virulence genotypes. *J Infect Dis***203**: 1844-1849.
515 Manges AR, Johnson JR, Foxman B, O'Bryan TT, Fullerton KE & Riley LW (2001)
516 Widespread distribution of urinary tract infections caused by a multidrug-resistant
517 *Escherichia coli* clonal group. *N Engl J Med***345**: 1007-1013.
518 Messika J, Magdoud F, Clermont O, Margetis D, Gaudry S, Roux D, Branger C, Dreyfuss D,
519 Denamur E & Ricard JD (2012) Pathophysiology of *Escherichia coli* ventilator-associated
520 pneumonia: implication of highly virulent extraintestinal pathogenic *Escherichia coli* strains.
521 *Intensive Care Med***38**: 2007-2016.
522 Moissenet D, Salauze B, Clermont O, Bingen E, Arlet G, Denamur E, Merens A, Mitanchez
523 D & Vu-Thien H (2010) Meningitis caused by *Escherichia coli* producing TEM-52 extended-
524 spectrum beta-lactamase within an extensive outbreak in a neonatal ward: epidemiological
525 investigation and characterization of the strain. *J Clin Microbiol***48**: 2459-2463.
526 Nicolas-Chanoine MH, Bertrand X & Madec JY (2014) *Escherichia coli* ST131, an intriguing
527 clonal group. *Clin Microbiol Rev***27**: 543-574.
528 Nicolas-Chanoine MH, Gruson C, Bialek-Davenet S, *et al.* (2013) 10-Fold increase (2006-11)
529 in the rate of healthy subjects with extended-spectrum beta-lactamase-producing *Escherichia*
530 *coli* faecal carriage in a Parisian check-up centre. *J Antimicrob Chemother***68**: 562-568.
531 Nowrouzian FL, Wold AE & Adlerberth I (2005) *Escherichia coli* strains belonging to
532 phylogenetic group B2 have superior capacity to persist in the intestinal microflora of infants.
533 *J Infect Dis***191**: 1078-1083.
534 Partridge SR, Tsafnat G, Coiera E & Iredell JR (2009) Gene cassettes and cassette arrays in
535 mobile resistance integrons. *FEMS Microbiol Rev***33**: 757-784.

Russo TA & Johnson JR (2003) Medical and economic impact of extraintestinal infections due to *Escherichia coli*: focus on an increasingly important endemic problem. *Microbes Infect***5**: 449-456.

Schubert S, Darlu P, Clermont O, Wieser A, Magistro G, Hoffmann C, Weinert K, Tenaillon O, Matic I & Denamur E (2009) Role of intraspecies recombination in the spread of pathogenicity islands within the *Escherichia coli* species. *PLoS Pathog***5**: e1000257.

Smati M, Clermont O, Le Gal F, Schichmanoff O, Jaureguy F, Eddi A, Denamur E & Picard B (2013) Real-time PCR for quantitative analysis of human commensal *Escherichia coli* populations reveals a high frequency of subdominant phylogroups. *Appl Environ Microbiol***79**: 5005-5012.

Tapader R, Chatterjee S, Singh AK, Dayma P, Haldar S, Pal A & Basu S (2014) The high prevalence of serine protease autotransporters of *Enterobacteriaceae* (SPATEs) in *Escherichia coli* causing neonatal septicemia. *Eur J Clin Microbiol Infect Dis***33**: 2015-2024.

Tenaillon O, Skurnik D, Picard B & Denamur E (2010) The population genetics of commensal *Escherichia coli*. *Nat Rev Microbiol***8**: 207-217.

Unno T, Han D, Jang J, Lee SN, Ko G, Choi HY, Kim JH, Sadowsky MJ & Hur HG (2009) Absence of *Escherichia coli* phylogenetic group B2 strains in humans and domesticated animals from Jeonnam Province, Republic of Korea. *Appl Environ Microbiol***75**: 5659-5666.

Walk ST, Alm EW, Gordon DM, Ram JL, Toranzos GA, Tiedje JM & Whittam TS (2009) Cryptic lineages of the genus *Escherichia*. *Appl Environ Microbiol***75**: 6534-6544.

Woerther PL, Angebault C, Jacquier H, *et al.*(2013) Characterization of fecal extended-spectrum-beta-lactamase-producing *Escherichia coli* in a remote community during a long time period. *Antimicrob Agents Chemother***57**: 5060-5066.

Table 1. Phylogenic, virulence and resistance characteristics of the 279 *E. coli* strains of the COLIVILLE collection

Phylogenetic groups/subgroups		Strains N (%)	Mean of virulence score* (± SD†)	Resistant strains‡ N (%)	Mean resistance score§ (± SD)
A		80 (28.7)	2.5 (± 2.5)	28 (35.0)	0.9 (± 1.5)
B1		36 (12.9)	2.3 (± 2.1)	17 (47.2)	1.0 (± 1.4)
B2		95 (34.0)	10.4 (± 3.0)	43 (45.3)	0.9 (± 1.3)
I‖	ST131	4 (4.2#)	7.8 (± 1.3)	2 (50.0**)	1.3 (± 1.5)
	Non ST131	9 (9.5)	7.2 (± 1.9)	8 (88.9)	1.6 (± 1.7)
II		21 (22.1)	12.5 (± 2.3)	8 (38.1)	0.6 (± 1.0)
III		2 (2.1)	11.5 (± 0.7)	1 (50.0)	0.5 (± 0.7)
IV		20 (21.1)	10.7 (± 3.0)	7 (35.0)	0.6 (± 1.1)
V		1 (1.1)	13.0	1 (100.0)	3.0
VI		3 (3.2)	13.7 (± 0.6)	2 (66.7)	1.7 (± 2.1)
VII		6 (6.3)	10.0 (± 1.7)	3 (50.0)	0.5 (± 0.6)
VIII		5 (5.2)	5.8 (± 1.3)	4 (80.0)	1.0 (± 0.7)
IX		16 (16.8)	11.2 (± 1.3)	4 (25.0)	0.4 (± 0.9)
X		1 (1.1)	4.0	1 (100.0)	2.0
Unassigned		7 (7.3)	10.0 (± 3.7)	4 (57.1)	1.7 (± 2.0)
C		9 (3.2)	5.2 (± 2.0)	8 (88.9)	2.9 (± 1.8)
D		25 (9.0)	4.9 (± 3.2)	14 (56.0)	1.7 (± 2.0)
CGA††		12 (48.0)	5.7 (± 3.2)	10 (83.3)	2.7 (± 1.7)
Non CGA		13 (52.0)	4.2 (± 3.1)	4 (30.8)	0.9 (± 1.8)
E		7 (2.5)	3.6 (± 1.9)	2 (28.6)	0.6 (± 1.1)
F		27 (9.7)	6.9 (± 3.8)	13 (48.1)	1.2 (± 1.6)
Total		279 (100.0)	5.9 (± 4.5)	125 (44.8)	1.1 (± 1.5)

* Virulence score, defined as the number of virulence genes present in the strain, among the 20 tested.
† SD, standard-deviation.
‡ Resistant strain, defined as a strain resistant to at least one antimicrobial agent amongthe 18 tested.
§ Resistance score, defined as the number of antibiotic classes (penicillins, cephalosporins, carbapenems, aminoglycosides, amphenicols, sulfonamides, tetracyclines, quinolones, furans, phosphonic acids) to which the strain is resistant amongthe 10 tested.
‖ The subgroups I, II, III, IV, V, VI, VII, VIII, IX and X correspond to the STc131, 73, 127, 141, 144, 12, 14, 452, 95 and 372, respectively (Clermont *et al.*, 2015). ST131, Sequence Type 131.
Proportions of subgroups are reported as fractions of the respective phylogroups.

** Proportions of resistant strains within the subgroup.
†† CGA, clonal group A (STc69).

595 **Table 2.** Virulence traits of the 279 *E. coli* strains of the COLIVILLE collection, with respect to the phylogenetic group of the strains

Virulence traits	Function	Strains N (%)	Phylogenetic groups							Global adjusted p values	Adjusted p values B2 versus other groups
			A (n= 80)	B1 (n= 36)	B2 (n= 95)	C (n= 9)	D (n= 25)	E (n= 7)	F (n= 27)		
iha	Adhesin	94 (33.7)	18	7	35	1	12	5	16	0.0004	0.4
papC	Adhesin	65 (23.3)	4	1	42	1	4	0	13	$9x10^{-12}$	$2x10^{-8}$
hra	Adhesin	61 (21.9)	11	5	36	2	6	1	0	$7x10^{-5}$	$9x10^{-6}$
sfa/foc	Adhesin	51 (18.3)	1	1	46	0	0	3	0	$6x10^{-16}$	$5x10^{-16}$
*papGII**	Adhesin	32 (11.5)	0	0	22	0	4	0	6	$6x10^{-7}$	$5x10^{-5}$
ibeA	Adhesin	30 (10.8)	0	0	29	0	1	0	0	$2x10^{-10}$	$5x10^{-14}$
*papGIII**	Adhesin	21 (7.5)	0	0	21	0	0	0	0	$4x10^{-7}$	$5x10^{-11}$
irp2	Iron acquisition	170 (60.9)	28	10	88	8	15	0	21	$6x10^{-16}$	$5x10^{-16}$
fyuA	Iron acquisition	168 (60.2)	27	9	88	8	15	0	21	$6x10^{-16}$	$5x10^{-16}$
iucC	Iron acquisition	128 (45.9)	26	7	58	4	14	5	14	$6x10^{-5}$	$5x10^{-4}$
iroN	Iron acquisition	95 (34.1)	5	7	71	4	3	3	2	$6x10^{-16}$	$5x10^{-16}$
ireA	Iron acquisition	57 (20.4)	12	3	30	4	2	2	4	0.005	0.002
ompT	Protectin	163 (58.4)	12	14	88	8	15	5	21	$6x10^{-16}$	$5x10^{-16}$
traT	Protectin	139 (49.8)	27	19	51	6	16	1	19	0.002	0.4
neuC	Protectin	66 (23.7)	9	0	41	0	1	0	15	$2x10^{-12}$	10^{-7}
usp	Toxin	106 (38.0)	1	0	90	0	0	0	15	$6x10^{-16}$	$5x10^{-16}$
sat	Toxin	72 (25.8)	15	1	25	1	12	0	18	$5x10^{-8}$	0.9
clbQ	Toxin	55 (19.7)	0	0	55	0	0	0	0	$6x10^{-16}$	$5x10^{-16}$
hlyC	Toxin	41 (14.7)	0	0	39	0	2	0	0	$2x10^{-15}$	$5x10^{-16}$
cnf1	Toxin	33 (11.8)	0	0	33	0	0	0	0	$5x10^{-13}$	$5x10^{-16}$

596 * Distinct alleles of the *papG* virulence gene.
597

Table 3. Distribution of resistance to 18 antibiotics in the COLIVILLE collection (n= 279)

Antibiotic classes	Antibiotics	Strains N (%)
Penicillins	Amoxicillin	83 (29.7)
	Amoxicillin – clavulanic acid	45 (16.1)
Cephalosporins	Cefoxitin	3 (1.1)
	Cefotaxime	1 (0.4)
Carbapenems	Ertapenem	0 (0.0)
	Imipenem	0 (0.0)
Aminoglycosides	Streptomycin	65 (23.3)
	Gentamicin	2 (0.7)
	Tobramycin	2 (0.7)
	Netilmicin	1 (0.4)
	Amikacin	0 (0.0)
Tetracyclines	Tetracycline	61 (21.9)
Sulfonamides	Cotrimoxazole	44 (15.8)
Amphenicols	Chloramphenicol	20 (7.2)
Quinolones	Nalidixic acid	17 (6.1)
	Ofloxacin	9 (3.2)
Phosphonic acids	Fosfomycin	1 (0.4)
Furans	Nitrofurantoin	1 (0.4)

FIGURE LEGENDS

Figure 1: Phenotypic and genotypic characteristics of *E. coli* strains ineach of the three collections.The COLIVILLE collection is the population isolated in (and labelled)"2010". The VDG ("1980") and AEM ("2000") collections are fromDuriez *et al.*(2001) and Escobar-Paramo *et al.* (2004), respectively.(a) Results are presented as a proportional stacked bar graph representing thephylogenetic distribution of*E. coli*. (b)Distribution of the 'reduced'virulence score. The black bars within each box plot show median values. The box covers the 25^{th} percentile to the 75^{th} percentile of the data. Bars above and below the box show 1.5 times the inter-quartile range. Dots located at some distance outside the bars correspond to outliers. (c)Histograms representingprevalence ofresistance to amoxicillin (AMX), amoxicillin-clavulanic acid (AMC), streptomycin (S), chloramphenicol (C), cotrimoxazole (SXT), tetracycline(TE) and nalidixic acid (NAL), and antibiotic resistant strains.

Fig. 1a

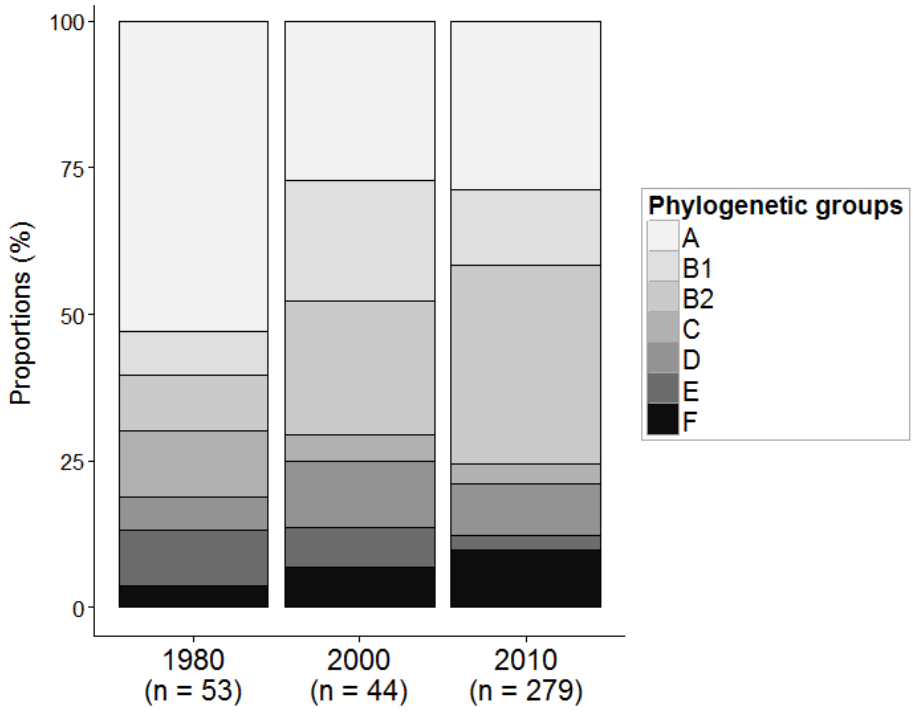
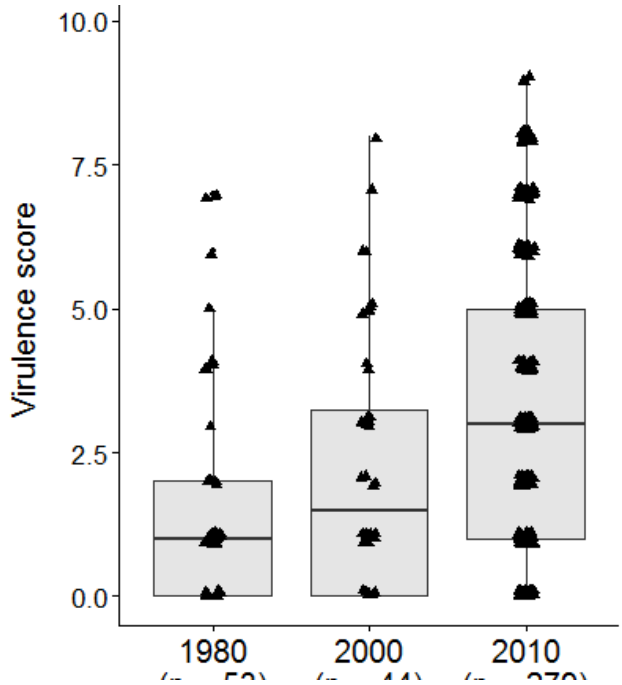
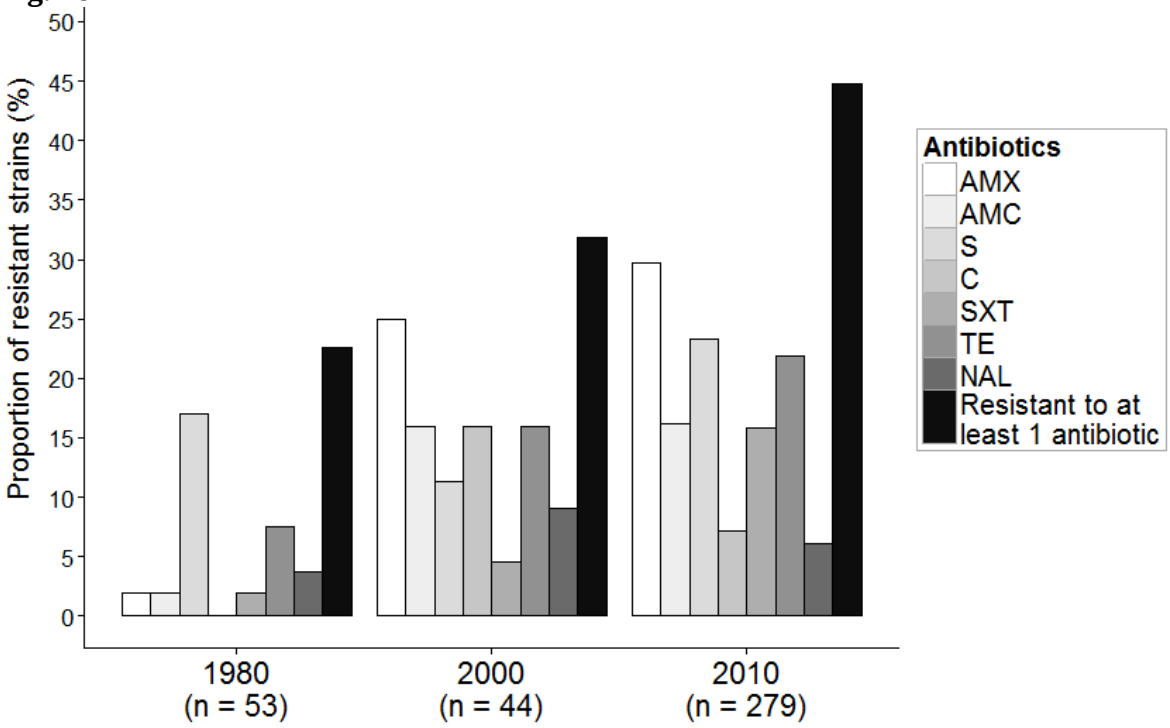

100
75
50
25
0
Proportions (%)
1980
(n = 53)
2000
(n = 44)
2010
(n = 279)
Phylogenetic groups
A
B1
B2
C
D
E
F

Fig. 1b


10.0
7.5
5.0
2.5
0.0
Virulence score
1980
2000
2010

668 **Fig. 1c**



669
670