Microbiology (2012), **158**, 2997–3004 **DOI** 10.1099/mic.0.060814-0

Antibiotic resistance plasmids spread among natural isolates of *Escherichia coli* in spite of CRISPR elements

Marie Touchon,[1,2] Sophie Charpentier,[3,4,5] Dominique Pognard,[3] Bertrand Picard,[4,6] Guillaume Arlet,[7,8] Eduardo P. C. Rocha,[1,2] Erick Denamur[4,5] and Catherine Branger[3,4,5]

Correspondence
Catherine Branger
catherine.branger@lmr.aphp.fr

[1]Institut Pasteur, Microbial Evolutionary Genomics, Département Génomes et Génétique, F-75015 Paris, France

[2]CNRS, UMR3525, 25 R. Dr Roux, F-75015 Paris, France

[3]AP-HP, Hôpital Louis Mourier, Service de Microbiologie-Hygiène, Colombes F-92701, France

[4]INSERM, UMR-S 722, Paris, F-75018, France

[5]Univ Paris Diderot, Sorbonne Paris Cité, UMR-S 722, Faculté de Médecine, Site Xavier Bichat, Paris, F-75018, France

[6]Univ Paris Nord, Sorbonne Paris Cité, UMR-S 722, Faculté de Médecine, Site Xavier Bichat, Paris, F-75018, France

[7]Univ Pierre et Marie Curie, Faculté de Médecine, Site Saint-Antoine, Laboratoire de Bactériologie, ER8, Paris, F-75012, France

[8]AP-HP, Hôpital Tenon, Service de Bactériologie-Hygiène, Paris, F-75020, France

Clustered, regularly interspaced, short palindromic repeats (CRISPRs) are implicated in defence against foreign DNA in various archaeal and bacterial species. They have also been associated with a slower spread of antibiotic resistance. However, experimental and evolutionary studies raise doubts about the role of CRISPRs as a sort of immune system in *Escherichia coli*. We studied a collection of 263 natural *E. coli* isolates from human and animal hosts, representative of the phylogenetic and lifestyle diversity of the species and exhibiting various levels of plasmid-encoded antibiotic resistance. We characterized the strains in terms of CRISPRs, performed replicon typing of the plasmids and tested for class 1 integrons to explore the possible association between CRISPRs and the absence of plasmids and mobile antibiotic resistance determinants. We found no meaningful association between the presence/absence of the *cas* genes, reflecting the activity of the CRISPRs, and the presence of plasmids, integrons or antibiotic resistance. No CRISPR in the collection contained a spacer that matched an antibiotic resistance gene or element involved in antibiotic resistance gene mobilization, and 79.8 % (210/263) of the strains lacked spacers matching sequences in the 2282 plasmid genomes available. Hence, *E. coli* CRISPRs do not seem to be efficient barriers to the spread of plasmids and antibiotic resistance, consistent with what has been reported for phages, and contrary to reports concerning other species.

Received 24 May 2012
Revised 19 September 2012
Accepted 10 October 2012

INTRODUCTION

Clustered, regularly interspaced, short palindromic repeats (CRISPRs) and their CRISPR-associated (*cas*) genes have been identified in most archaea and many bacteria. CRISPRs typically consist of short (23–47 bp) and highly conserved direct repeats interspersed with non-repetitive sequences called spacers (Sorek *et al.*, 2008). The CRISPRs form peculiar genetic loci, which may provide acquired immunity against viruses and plasmids by targeting nucleic acids in a sequence-specific manner (Horvath & Barrangou, 2010; Wiedenheft *et al.*, 2012). CRISPR-encoded immunity is provided by transcription of the repeat-spacer array, followed by transcript processing into small CRISPR RNAs, which are then used in

Abbreviations: BHR, broad-host-range; CRISPR, clustered, regularly interspaced, short palindromic repeat; FAC, factorial analysis of correspondence.

Four supplementary tables are available with the online version of this paper.

combination with Cas proteins to interfere with invasive DNA or RNA (Sapranauskas *et al.*, 2011). However, the number, sequence and length of CRISPR arrays and the number and sequence of the Cas proteins are highly variable between bacterial species (Makarova *et al.*, 2011). In *Escherichia coli*, CRISPRs map in two largely independent pairs of loci, CRISPR1 and 2 located at 62 min on the chromosome, and CRISPR3 and 4 at 20 min, each including a different type of *cas* gene (subtype I-E and subtype I-F, respectively) (Diez-Villaseñor *et al.*, 2010; Touchon & Rocha, 2010). The functional roles of the CRISPRs in *E. coli* remain unclear. Effective CRISPR-based protection against phages and/or plasmids cannot be observed in wild-type strains, and its demonstration requires complex genetic manipulation such as H-NS inactivation (Edgar & Qimron, 2010; Pougach *et al.*, 2010; Pul *et al.*, 2010), LeuO overexpression (Westra *et al.*, 2010), expression of a plasmid-encoded protein targeted for export in cells that lack DnaK (Perez-Rodriguez *et al.*, 2011), or high-temperature protein G expression (Yosef *et al.*, 2011). It is doubtful that such conditions are relevant in the natural environment. Molecular phylogeny analysis of several hundred natural *E. coli* isolates has shown a strong phylogenetic inertia of CRISPR loci, with CRISPR being almost absent from strains of the major B2 phylogenetic group, and a very slow spacer composition evolution; these observations are at odds with expectations for an active immune system (Delannoy *et al.*, 2012; Diez-Villaseñor *et al.*, 2010; Touchon & Rocha, 2010; Touchon *et al.*, 2011). Lastly, *cas* genes evolve under purifying selection, without any evidence of positive selection, indicating that there is no evidence that they are involved in an arms race between bacterial and archaeal hosts and invading virus or plasmid DNA (Takeuchi *et al.*, 2012). These various observations led to the suggestion that in *E. coli*, because CRISPRs are maintained in most lineages, the CRISPR/Cas system may have functions in addition to defence against foreign DNA (Babu *et al.*, 2011).

The aim of this work was to gain further insights into the role of CRISPRs in *E. coli* by analysing a collection of 263 natural *E. coli* isolates from human and animal hosts, representative of the phylogenetic and lifestyle diversity of the species and exhibiting various levels of plasmid-encoded antibiotic resistance. In this collection, we characterized the CRISPR content and the presence of mobile elements, i.e. plasmids and integrons, and explored the putative link between the presence/absence of the CRISPRs/Cas systems, the presence of mobile elements and antibiotic resistance. We also searched for CRISPR spacers that target nucleic acid sequences in relation to antibiotic resistance and the acquisition or dissemination of plasmids and integrons.

METHODS

Bacterial strains and antibiotic susceptibility. The collection of 263 natural *E. coli* isolates used for this study has been described in Touchon *et al.* (2011). Briefly, they were isolated from various hosts [humans (n=217), domestic animals (n=22), wild animals (n=24)] over a 25-year period (1980–2005) (Branger *et al.*, 2005; Lefort *et al.*, 2011; Picard *et al.*, 1999; Skurnik *et al.*, 2006). The strains were either commensal (119 strains) or associated with extraintestinal infections (144 strains). The CRISPR sequences and the phylogroups (Moissenet *et al.*, 2010; Tenaillon *et al.*, 2010) as determined by specific PCR (Clermont *et al.*, 2000; Clermont *et al.*, 2011a) and multilocus sequence typing (Clermont *et al.*, 2011b; Jaureguy *et al.*, 2008) were available for all the strains (Touchon *et al.*, 2011). The proportions of the phylogroups in the collection are: A (23.5 %), B1 (17.1 %), B2 (33.8 %), C (5.5 %), D (12.5 %), F (4.9 %), *Escherichia* clades (0.7 %) and ungrouped (2.6 %). Extended-spectrum β-lactamases (ESBLs) were produced by 125 of the strains, including 53, 22 and 50 strains producing a TEM-, SHV- and CTX-M-type ESBL, respectively, all encoded by plasmid genes (Branger *et al.*, 2005; Marcadé *et al.*, 2009).

For this study, resistance (mainly plasmid-encoded) to seven antibiotics (ampicillin, cefotaxime, streptomycin, kanamycin, tetracycline, chloramphenicol and sulphonamide) was determined for all the strains. The strains were also tested for penicillinase by PCR or isoelectric focusing (Branger *et al.*, 1997, 2005): 32 non-ESBL- and 59 ESBL-producing strains were penicillinase-positive.

The main characteristics of the strains are given in Table S1 available with the online version of this paper.

Plasmid replicon typing and integron detection. Plasmid replicons were determined using the PCR-based replicon typing scheme, described by Carattoli *et al.* (2005), which detects 18 plasmid replicons frequently found in *Enterobacteriaceae*. This technique involves the detection of the coding sequence of *repA*, the *cis*-repeats of the origin of replication or the countertranscript RNAI (Carattoli *et al.*, 2005), all elements defining the basic replicons of plasmids.

The strains were screened for the presence of class 1 integrons, the most prevalent integrons in *E. coli* (Lefort *et al.*, 2011; Skurnik *et al.*, 2005), by PCR amplification of an internal fragment of the integrase gene with a previously described primer pair (intl1L and intl1R) (Ploy *et al.*, 2000).

CRISPR spacer analysis. To identify the spacer sequences matching sequences from plasmids (proto-spacers), the spacers were subjected to the standard BLASTN search (e-value threshold, 1.10^{-5}) against 2282 plasmid genomes available in GenBank. Proto-spacers were defined by identifying homologous plasmid sequences with an e-value $<1.10^{-5}$ and less than 10 % difference in sequence length.

Statistical analysis. The Pearson's χ^2 or the Fisher's exact test was used, when appropriate, for comparisons. The Student's t test was used for comparing means. A P value <0.05 was considered significant for all the tests. For correlation analyses, we used the Spearman rho correlation coefficient. Factorial analysis of correspondence (FAC) was used to describe associations between variables. FAC uses a covariance matrix based on χ^2 distances. The computation determines a plane defined by two principal axes of the analysis; the first axis, F1, accounts for most of the variance, and the second axis, F2, which is orthogonal to F1, accounts for the largest part of the variance not accounted for by F1 (Greenacre, 1992). SPAD.N software (Cisia) was used for FAC with a two-way table. The table had 263 rows, corresponding to the 263 *E. coli* strains studied, and 30 columns, corresponding to the 30 variables: the seven phylogenetic groups (A, B1, B2, D, C, F and the *Escherichia* clades), the ungrouped strains, the presence of ESBL and of penicillinase, the three resistance scores (absence of resistance, resistance to one to three antibiotics, resistance to four to seven antibiotics), the presence of class 1 integrons, the presence of five types of replicons [IncF (including IncFII, IncFIA, IncFIB and IncFIC), IncB/O, Inc11, broad-host-range (BHR) replicons including IncL/M, IncA/C, IncN and IncP, and other

replicons (IncK, IncHI1, IncHI2 and IncY)], the absence of described replicons, the presence of two types of *cas* gene (subtype I-E and subtype I-F), the absence of *cas* genes, the presence of the four CRISPR loci (CRISPR1–4), and the total number of repeats classified into three categories (2–9, 10–24 and 25–54). For each strain, each variable was coded as a binary code: present=1, absent=0.

RESULTS AND DISCUSSION

Characterization of the plasmid and integron content of the collection

We used a PCR-based replicon typing scheme to analyse the plasmid content of the strains (Carattoli *et al.*, 2005). Although some plasmids are not typable with this scheme (Carattoli, 2009; Marcadé *et al.*, 2009), the results of such PCR typing correlate well with those of S1 nuclease plasmid analysis in a collection of natural *E. coli* and can therefore be considered to be a good approximation of plasmid content (Bengtsson *et al.*, 2012).

Replicon typing detected 506 replicons in 221 strains (84%) (mean 1.92, range 0–6): 55 strains had one replicon, 78 strains had two, 63 strains had three, 21 strains had four, two strains had five and two strains had six. Fourteen different incompatibility groups were detected: IncF (n=186) [IncFII (n=155), IncFIB (n=142), IncFIA (n=48) and IncFIC (n=10)], IncI1 (n=37), IncB/O (n=14), BHR replicons (n=70) [IncA/C (n=46), IncN (n=15), IncP (n=9) and IncL/M (n=4)], other replicons (n=25) [IncY (n=9), IncK (n=8), IncHI1 (n=5) and IncHI2 (n=4)].

As expected, replicons were more prevalent among resistant strains than susceptible strains (89.2 % versus 72.7 % contained replicons, respectively; P=0.002). Among strains with plasmids, resistant strains had more different replicons than susceptible strains (mean 2.4 versus 1.92, respectively; P=0.003); similarly, more ESBL-producing strains than non-ESBL-producing strains carried replicons [(92.8 % versus 76.81 %, respectively; P=0.0003) (mean number of replicons 2.44 versus 2.12, respectively; P=0.02)]. There were no such differences between non-ESBL-producing strains producing and not producing penicillinase.

BHR replicons were detected only in resistant strains (70/186 strains) (P<0.0001), which were mostly ESBL-producing strains (n=61) (P<0.0001); most of the non-ESBL strains produced penicillinase (n=7) (P<0.0001). In the ESBL-producing strains, most of these replicons are on the same plasmid as the ESBL genes (Marcadé *et al.*, 2009). IncI1 replicons were found more frequently among ESBL-producing strains than among all other strains (19.2 % versus 10.1 %; P=0.036). However, combining all resistance markers into one group, there was no difference in the frequency of IncI1 replicons between resistant and susceptible strains. The carriage of the IncF or IncB/O replicons was not associated with the antibiotic resistance of the strains. The remaining replicons were generally more frequent in the ESBL-producing than other strains (P=0.029).

We screened the collection for class 1 integrons. Integrons are mobile elements associated with antibiotic resistance, and class 1 integrons are the most prevalent (Mazel, 2006). Integrons were found in 135 strains, all of which were resistant to antibiotics and most of which produced an ESBL (120 strains) (96 % of the ESBL-producing strains versus 10.8 % of the non-ESBL producing strains, P<0.0001); seven (21.8 %) of the non-ESBL-producing strains produced penicillinase and eight (7.5 %) did not (P=0.02).

Comparison of CRISPRs/*cas* genes according to the presence of replicons and integrons in *E. coli* strains

Cas proteins are necessary for the functioning of the CRISPR/Cas immune system and are indicators of the system's activity (Horvath & Barrangou, 2010). To test for an association between the presence and the activity of CRISPRs and the presence of replicons and integrons, we compared the presence of *cas* genes (both subtype I-E and I-F) between *E. coli* strains with and without replicons or integrons. There were no significant differences in the prevalence of *cas* genes between the replicon- or integron-positive and -negative *E. coli* strains (Table 1).

Plasmids of the IncF, IncB/O and IncI1 groups have narrow host ranges and are commonly found in *Enterobacteriaceae* (Johnson *et al.*, 2007); IncF plasmids are particularly well adapted to *E. coli* (Boyd *et al.*, 1996). Other groups of plasmids (IncL/M, IncA/C, IncN and IncP) are considered to have BHRs (Toukdarian, 2004). We analysed the data according to the type of replicon, because CRISPRs may be particularly involved in the defence against BHR replicons. We found no significant differences in the prevalence of the *cas* system except for strains containing IncF replicons, which had fewer *cas* genes (28.4 % versus 41.5 %; P=0.039) (Table 2). These data generally indicate that there is no significant link between CRISPR activity and the presence of foreign DNA (plasmids or integrons).

CRISPRs and acquired antibiotic resistance

Recently, an inverse correlation between the presence of a CRISPR/Cas locus in enterococci and that of acquired antibiotic resistance genes has been reported (Palmer & Gilmore, 2010). Thus, we compared CRISPRs between susceptible and resistant strains, assuming that the observed resistances were largely due to plasmid-borne determinants. We also looked for differences in the production of plasmid-encoded enzymes conferring resistance to β-lactams, i.e. penicillinase and ESBL.

The presence of subtype I-E *cas* genes did not differ between the susceptible and resistant strains, or between those producing and not producing β-lactamases. The subtype I-F *cas* system was significantly more prevalent among susceptible strains [n=11 (14.2 %) versus n=5 (2.6 %); P=0.0003] (Table 1). However, in our collection of strains this system was carried only by a small number of

Table 1. CRISPR characteristics of the *E. coli* strains studied according to antibiotic resistance, β-lactamase production and presence of integrons and replicons

	Antibiotic resistance			ESBL production			Non-ESBL production			Integron			Replicon		
	Presence	Absence	*P*	Presence	Absence	*P*	Presence of Pase*	Absence of Pase	*P*	Presence	Absence	*P*	Presence	Absence	*P*
No. of strains	186	77		125	138		32	106		135	128		222	41	
No. (%) of strains with:															
Subtype I-E	118 (63.4)	45 (58.4)	0.44	78 (62.4)	85 (61.5)	0.81	18 (56.2)	67 (63.2)	0.31	87 (64.4)	76 (59.3)	0.39	142 (63.9)	21 (51.2)	0.08
Subtype I-F	5 (2.6)	11 (14.2)	0.0003	5 (4.0)	11 (7.9)	0.17	0 (0)	11 (10.3)	0.057	5 (3.7)	11 (8.5)	0.09	13 (5.8)	3 (7.3)	0.71
Absence of *cas* gene	64 (34.4)	21 (27.2)	0.26	43 (34.4)	42 (30.4)	0.49	14 (43.7)	28 (26.4)	0.061	44 (32.5)	41 (32)	0.9	68 (30.6)	17 (41.4)	0.17

*Pase, Penicillinase.

Table 2. CRISPR characteristics of the *E. coli* strains studied according to the five different types of replicons

	IncF replicon			IncI1 replicon			IncB/O replicon			BHR replicons			Other replicons		
	Presence	Absence	*P*	Presence	Absence	*P*	Presence	Absence	*P*	Presence	Absence	*P*	Presence	Absence	*P*
No. of strains	186	77		37	226		14	249		70	193		25	238	
No. (%) of strains with:															
Subtype I-E	122 (65.5)	41 (53)	0.06	27 (73)	136 (60.1)	0.13	8 (57.1)	155 (62.2)	0.70	48 (68.5)	115 (59.5)	0.27	18 (72)	145 (61)	0.27
Subtype I-F	12 (6.4)	4 (5.1)	0.69	2 (5.4)	14 (6.2)	0.85	1 (7.1)	15 (6.0)	0.86	2 (2.8)	14 (7.2)	0.18	0 (0)	16 (6.7)	0.18
Absence of *cas* gene	53 (28.4)	32 (41.5)	0.039	9 (24.3)	76 (33.6)	0.26	5 (35.7)	80 (32.1)	0.78	20 (28.6)	65 (33.6)	0.43	7 (28)	78 (32.7)	0.62

strains of the B2 group (15/89) and a single B1 strain: most group B2 strains with subtype I-F systems were susceptible to antibiotics (11/15).

Comparisons excluding the phylogenetic group B2 strains

The major phylogenetic group B2 is almost devoid of CRISPR (Diez-Villaseñor *et al.*, 2010; Touchon *et al.*, 2011). This may bias the analysis of our collection, so we reanalysed the data, excluding the 89 group B2 strains (19 replicon-negative and 70 replicon-positive strains): almost all (92.5 %) of the 174 non-B2 group strains possessed *cas* genes.

The distributions of replicons, integrons, antibiotic resistance and β-lactamases in the 174 non-B2 group strains were similar to those in the total collection: 152 (87.3 % versus 84.4 %, respectively) had at least one replicon, 92 had an integron (52.8 % versus 51.3 %, respectively), 48 (27.5 % versus 29.2 %, respectively) were antibiotic susceptible, 83 (47.7 % versus 47.5 %, respectively) produced an ESBL, and among the 91 non-ESBL-producing strains, 20 (21.9 % versus 23.1 %, respectively) produced a penicillinase.

The findings for the non-B2 group strains were similar to those for the entire collection including group B2 strains: there was no correlation between the presence of the subtype I-E *cas* system and any characteristics of the strains including the type of replicons. The subtype I-F *cas* system is not found in non-B2 strains, with one exception, as previously reported (Diez-Villaseñor *et al.*, 2010; Touchon & Rocha, 2010). Thus, exclusion of B2 strains did not substantially change the results.

Numbers of repeats according to the presence of replicons, integrons or antibiotic resistance

Another indicator of CRISPR activity could be the median number of repeats. Repeat-spacer units can be both added and deleted, such that a CRISPR can evolve while maintaining an identical number of repeats, although species with a large number of repeats nevertheless tend to have high CRISPR activity (Delaney *et al.*, 2012; Kuno *et al.*, 2012; Tyson & Banfield, 2008). Using this measure, we found no evidence of any negative association between the presence of replicons, integrons or antibiotic resistance and the total number of repeats (data not shown).

Multidimensional analysis of the data

We conducted an FAC to have an overview of the data. The first two axes (planes F1 and F2) accounted for 33.56 % of the total variance. The first factor separated the presence of CRISPR1 and 2, large numbers of CRISPR repeats (10–24 and 25–54 categories) and the presence of subtype I-E *cas* genes from small numbers of CRISPR repeats (two to nine category) and the absence of subtype I-E *cas* genes (Fig. 1). Thus, factor F1, which accounts for 19.57 % of the total variance, may be associated with the activity of the CRISPR system. The phylogenetic groups were projected along this axis, and group B2 appeared on the opposite side from the others, illustrating the association between the species phylogeny and CRISPRs (Diez-Villaseñor *et al.*, 2010; Touchon *et al.*, 2011). The second factor, F2, accounted for 13.99 % of the total variance; it appeared to characterize the resistance and the plasmid content of the strains. Thus, a high resistance score (resistance to four to seven antibiotics) and the presence of a class 1 integron and an ESBL were projected at opposite extremes to the absence of resistance, a low level of resistance (resistance to one to three antibiotics) and the presence of a penicillinase. The presence and types of replicons were projected along this axis, with BHR, the other replicons, the IncI1 replicons and the presence of replicons being opposed to the presence of IncB/O replicons and the absence of replicons. The subtype I-F *cas* genes and the CRISPR4 were closely linked, as previously reported (Diez-Villaseñor *et al.*, 2010; Touchon *et al.*, 2011), and distinguished by the negative values of the two axes.

This FAC indicates that the strain phylogeny and the CRISPR activity are correlated (axis F1) and that the resistance to antibiotics and the presence of replicons and integrons are also correlated (axis F2), but that strain phylogeny and CRISPRs are independent of antibiotic resistance and the presence of integrons and replicons.

Sequence similarities between the spacers and available plasmid sequences

We searched for proto-spacers in 2282 sequenced prokaryote plasmids to assess the role of CRISPRs in the acquisition of plasmids. Most (210/263) strains lacked spacers that matched plasmid sequences. In 53 strains, we identified only 29 different spacers with significant matches (19 spacers had 100 % and 10 had 90.6–96.8 % identity to plasmid sequences); of these, 17 spacers matched sequences in genes encoding functions such as methylase, partition proteins, antirestriction proteins, and replication and transfer systems (Table S2). We found no occurrence of spacers that matched antibiotic resistance genes or elements involved in antibiotic resistance gene mobilization (e.g. Tn*3*, IS*Ecp1*, *intI*). We observed no association between spacers that matched plasmid sequences and ESBL production or the type of ESBL or replicon. The number of spacers matching plasmid genes in the 53 strains was relatively small: 77.3 % of the strains had one, 13.2 % had two, 7.5 % had three or four and one single strain (IAI48) had 15 (Tables S2, S3 and S4). These spacers were specific for each group of CRISPR. In CRISPR1 and CRISPR2, their diversity was low: only seven different spacers were found in the CRISPR1 of 27 strains and three in the CRISPR2 of 13 strains. For the CRISPR3–4/subtype I-F *cas* system, the diversity of the spacers was greater: eight different spacers were found in the CRISPR3 of 14 strains and 11 in CRISPR4 of seven

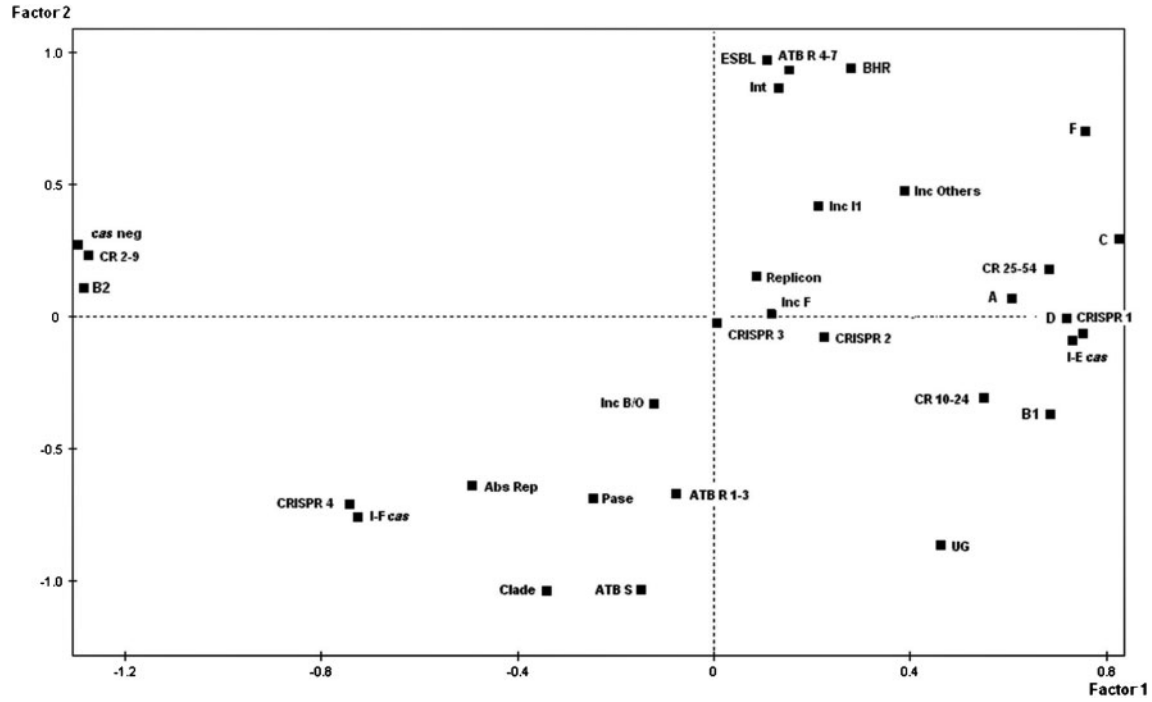

Fig. 1. Projection of the variables onto the planes defined by the first two axes of the FAC. Variables characterizing the 263 *E. coli* strains included in the analysis are: the *E. coli* phylogenetic group (A, B1, B2, C, D or F), the *Escherichia* clade (Clade) and the ungrouped strains (UG), the antibiotic resistance scores [absence of resistance (ATB S), resistance to one to three antibiotics (ATB R 1–3), resistance to four to seven antibiotics (ATB R 4–7)], the presence of penicillinase (Pase), ESBL and class 1 integron (Int), the five types of replicon [IncF, IncB/O, IncI1, BHR, other replicons (Inc Others)], the presence of at least one replicon (Replicon), the absence of replicons (Abs Rep), the presence of the four CRISPR loci (CRISPR1, 2, 3, 4), the numbers of CRISPR repeats in three categories: [two to nine (CR 2–9), 10 to 24 (CR 10–24) and 25 to 54 (CR 25–54)], the presence of two types of *cas* genes [subtype I-E (I-E *cas*) and subtype I-F (I-F *cas*)] and the absence of the *cas* gene (*cas* neg).

strains. Strains with the same spacers matching plasmids belonged to the same CRISPR spacer relatedness group as defined by Touchon *et al.* (2011), reflecting phylogenetic inertia (Tables S3 and S4).

Concluding remarks

We were not able to find any solid or meaningful correlation between the presence of plasmids, integrons, acquired antibiotic resistance and CRISPRs. This is surprising given the suggested role of CRISPRs in moderating plasmid transfer (Horvath & Barrangou, 2010); it is also contrary to findings in enterococci, in which CRISPR are negatively associated with antibiotic resistance (Palmer & Gilmore, 2010). Indeed, the probability of acquiring plasmids and integrons and thereby becoming multi-resistant should be higher for strains with less active CRISPRs, if such activity inhibits transfer. Our findings are consistent with the observation that the sensitivity of an *E. coli* collection, representative of the species diversity, to a set of 59 coliphages (Kutter, 2009) does not correlate with CRISPR content (Diez-Villaseñor *et al.*, 2010). The patterns of distribution that we report suggest that the presence of CRISPRs has little effect on the epidemiology of plasmids in *E. coli* or on the spread of antibiotic resistance. The 20 % of the strains that retain spacers with identity to plasmids could indicate ancient contacts with plasmids, the CRISPRs being no longer active.

ACKNOWLEDGEMENTS

This work was partly funded by a grant from the Assistance Publique-Hôpitaux de Paris (Contrat d'Initiation à la Recherche Clinique 05 103) and by the Fondation pour la Recherche Médicale. M. T and E. P. C. R. are supported by the CNRS and the Institut Pasteur.

REFERENCES

Babu, M., Beloglazova, N., Flick, R., Graham, C., Skarina, T., Nocek, B., Gagarinova, A., Pogoutse, O., Brown, G. & other authors (2011). A dual function of the CRISPR–Cas system in bacterial antivirus immunity and DNA repair. *Mol Microbiol* **79**, 484–502.

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.

Boyd, E. F., Hill, C. W., Rich, S. M. & Hartl, D. L. (1996). Mosaic structure of plasmids from natural populations of *Escherichia coli*. *Genetics* **143**, 1091–1100.

Branger, C., Bruneau, B., Lesimple, A. L., Bouvet, P. J., Berry, P., Sevali-Garcia, J. & Lambert-Zechovsky, N. (1997). Epidemiological typing of extended-spectrum β-lactamase-producing *Klebsiella pneumoniae* isolates responsible for five outbreaks in a university hospital. *J Hosp Infect* **36**, 23–36.

Branger, C., Zamfir, O., Geoffroy, S., Laurans, G., Arlet, G., Thien, H. V., Gouriou, S., Picard, B. & Denamur, E. (2005). Genetic background of *Escherichia coli* and extended-spectrum β-lactamase type. *Emerg Infect Dis* **11**, 54–61.

Carattoli, A. (2009). Resistance plasmid families in *Enterobacteriaceae*. *Antimicrob Agents Chemother* **53**, 2227–2238.

Carattoli, A., Bertini, A., Villa, L., Falbo, V., Hopkins, K. L. & Threlfall, E. J. (2005). Identification of plasmids by PCR-based replicon typing. *J Microbiol Methods* **63**, 219–228.

Clermont, O., Bonacorsi, S. & Bingen, E. (2000). Rapid and simple determination of the *Escherichia coli* phylogenetic group. *Appl Environ Microbiol* **66**, 4555–4558.

Clermont, O., Gordon, D. M., Brisse, S., Walk, S. T. & Denamur, E. (2011a). Characterization of the cryptic *Escherichia* lineages: rapid identification and prevalence. *Environ Microbiol* **13**, 2468–2477.

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

Delaney, N. F., Balenger, S., Bonneaud, C., Marx, C. J., Hill, G. E., Ferguson-Noel, N., Tsai, P., Rodrigo, A. & Edwards, S. V. (2012). Ultrafast evolution and loss of CRISPRs following a host shift in a novel wildlife pathogen, *Mycoplasma gallisepticum*. *PLoS Genet* **8**, e1002511.

Delannoy, S., Beutin, L., Burgos, Y. & Fach, P. (2012). Specific detection of enteroaggregative hemorrhagic *Escherichia coli* O104 : H4 strains using the CRISPR locus as target for a diagnostic real-time PCR. *J Clin Microbiol* **50**, 3485–3492.

Díez-Villaseñor, C., Almendros, C., García-Martínez, J. & Mojica, F. J. (2010). Diversity of CRISPR loci in *Escherichia coli*. *Microbiology* **156**, 1351–1361.

Edgar, R. & Qimron, U. (2010). The *Escherichia coli* CRISPR system protects from λ lysogenization, lysogens, and prophage induction. *J Bacteriol* **192**, 6291–6294.

Greenacre, M. (1992). Correspondence analysis in medical research. *Stat Methods Med Res* **1**, 97–117.

Horvath, P. & Barrangou, R. (2010). CRISPR/Cas, the immune system of bacteria and archaea. *Science* **327**, 167–170.

Jaureguy, F., Landraud, L., Passet, V., Diancourt, L., Frapy, E., Guigon, G., Carbonnelle, E., Lortholary, O., Clermont, O. & other authors (2008). Phylogenetic and genomic diversity of human bacteremic *Escherichia coli* strains. *BMC Genomics* **9**, 560.

Johnson, T. J., Wannemuehler, Y. M., Johnson, S. J., Logue, C. M., White, D. G., Doetkott, C. & Nolan, L. K. (2007). Plasmid replicon typing of commensal and pathogenic *Escherichia coli* isolates. *Appl Environ Microbiol* **73**, 1976–1983.

Kuno, S., Yoshida, T., Kaneko, T. & Sako, Y. (2012). Intricate interactions between the bloom-forming cyanobacterium *Microcystis aeruginosa* and foreign genetic elements, revealed by diversified clustered regularly interspaced short palindromic repeat (CRISPR) signatures. *Appl Environ Microbiol* **78**, 5353–5360.

Kutter, E. (2009). Phage host range and efficiency of plating. *Methods Mol Biol* **501**, 141–149.

Lefort, A., Panhard, X., Clermont, O., Woerther, P. L., Branger, C., Mentré, F., Fantin, B., Wolff, M., Denamur, E. & COLIBAFI Group (2011). Host factors and portal of entry outweigh bacterial determinants to predict the severity of *Escherichia coli* bacteremia. *J Clin Microbiol* **49**, 777–783.

Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F. J., Wolf, Y. I. & other authors (2011). Evolution and classification of the CRISPR–Cas systems. *Nat Rev Microbiol* **9**, 467–477.

Marcadé, G., Deschamps, C., Boyd, A., Gautier, V., Picard, B., Branger, C., Denamur, E. & Arlet, G. (2009). Replicon typing of plasmids in *Escherichia coli* producing extended-spectrum β-lactamases. *J Antimicrob Chemother* **63**, 67–71.

Mazel, D. (2006). Integrons: agents of bacterial evolution. *Nat Rev Microbiol* **4**, 608–620.

Moissenet, D., Salauze, B., Clermont, O., Bingen, E., Arlet, G., Denamur, E., Mérens, A., Mitanchez, D. & Vu-Thien, H. (2010). Meningitis caused by *Escherichia coli* producing TEM-52 extended-spectrum β-lactamase within an extensive outbreak in a neonatal ward: epidemiological investigation and characterization of the strain. *J Clin Microbiol* **48**, 2459–2463.

Palmer, K. L. & Gilmore, M. S. (2010). Multidrug-resistant enterococci lack CRISPR-*cas*. *MBio* **1**, e00227-10.

Perez-Rodriguez, R., Haitjema, C., Huang, Q., Nam, K. H., Bernardis, S., Ke, A. & DeLisa, M. P. (2011). Envelope stress is a trigger of CRISPR RNA-mediated DNA silencing in *Escherichia coli*. *Mol Microbiol* **79**, 584–599.

Picard, B., Garcia, J. S., 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.

Ploy, M. C., Denis, F., Courvalin, P. & Lambert, T. (2000). Molecular characterization of integrons in *Acinetobacter baumannii*: description of a hybrid class 2 integron. *Antimicrob Agents Chemother* **44**, 2684–2688.

Pougach, K., Semenova, E., Bogdanova, E., Datsenko, K. A., Djordjevic, M., Wanner, B. L. & Severinov, K. (2010). Transcription, processing and function of CRISPR cassettes in *Escherichia coli*. *Mol Microbiol* **77**, 1367–1379.

Pul, Ü., Wurm, R., Arslan, Z., Geißen, R., Hofmann, N. & Wagner, R. (2010). Identification and characterization of *E. coli* CRISPR-*cas* promoters and their silencing by H-NS. *Mol Microbiol* **75**, 1495–1512.

Sapranauskas, R., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P. & Siksnys, V. (2011). The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res* **39**, 9275–9282.

Skurnik, D., Le Menac'h, A., Zurakowski, D., Mazel, D., Courvalin, P., Denamur, E., Andremont, A. & Ruimy, R. (2005). Integron-associated antibiotic resistance and phylogenetic grouping of *Escherichia coli* isolates from healthy subjects free of recent antibiotic exposure. *Antimicrob Agents Chemother* **49**, 3062–3065.

Skurnik, D., Ruimy, R., Andremont, A., Amorin, C., Rouquet, P., Picard, B. & Denamur, E. (2006). Effect of human vicinity on antimicrobial resistance and integrons in animal faecal *Escherichia coli*. *J Antimicrob Chemother* **57**, 1215–1219.

Sorek, R., Kunin, V. & Hugenholtz, P. (2008). CRISPR – a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nat Rev Microbiol* **6**, 181–186.

Takeuchi, N., Wolf, Y. I., Makarova, K. S. & Koonin, E. V. (2012). Nature and intensity of selection pressure on CRISPR-associated genes. *J Bacteriol* **194**, 1216–1225.

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

Touchon, M. & Rocha, E. P. (2010). The small, slow and specialized CRISPR and anti-CRISPR of *Escherichia* and *Salmonella*. *PLoS ONE* **5**, e11126.

Touchon, M., Charpentier, S., Clermont, O., Rocha, E. P., Denamur, E. & Branger, C. (2011). CRISPR distribution within the *Escherichia coli* species is not suggestive of immunity-associated diversifying selection. *J Bacteriol* **193**, 2460–2467.

Toukdarian, A. (2004). Plasmid strategies for broad-host-range replication in Gram-negative bacteria. In *Plasmid Biology*, pp. 259–270. Edited by B. E. Funnell & G. J. Phillips. Washington, DC: American Society for Microbiology.

Tyson, G. W. & Banfield, J. F. (2008). Rapidly evolving CRISPRs implicated in acquired resistance of microorganisms to viruses. *Environ Microbiol* **10**, 200–207.

Westra, E. R., Pul, U., Heidrich, N., Jore, M. M., Lundgren, M., Stratmann, T., Wurm, R., Raine, A., Mescher, M. & other authors (2010). H-NS-mediated repression of CRISPR-based immunity in *Escherichia coli* K12 can be relieved by the transcription activator LeuO. *Mol Microbiol* **77**, 1380–1393.

Wiedenheft, B., Sternberg, S. H. & Doudna, J. A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. *Nature* **482**, 331–338.

Yosef, I., Goren, M. G., Kiro, R., Edgar, R. & Qimron, U. (2011). High-temperature protein G is essential for activity of the *Escherichia coli* clustered regularly interspaced short palindromic repeats (CRISPR)/Cas system. *Proc Natl Acad Sci U S A* **108**, 20136–20141.

Edited by: E. M. Top