Infection, Genetics and Evolution 74 (2019) 103916

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research paper

Polymorphism of Type I–F CRISPR/Cas system in *Escherichia coli* of phylogenetic group B2 and its application in genotyping

Jinzhao Long[a], Yake Xu[a,c], Liuyang Ou[a], Haiyan Yang[a], Yuanlin Xi[a], Shuaiyin Chen[a,*], Guangcai Duan[a,b,*]

[a] College of Public Health, Zhengzhou University, Zhengzhou, Henan, China
[b] Henan Innovation Center of Molecular Diagnosis and Laboratory Medicine, Xinxiang Medical University, Xinxiang, Henan, China
[c] School Hospital, Henan University of Science and Technology, Luoyang, Henan, China

ARTICLE INFO

Keywords:
CRISPR/Cas system
Type I–F
Genotyping
E. coli
MLST
Phylogroup B2

ABSTRACT

E. coli of phylogenetic group B2 is responsible for many extraintestinal infections, posing a great threat to health. The relatively polymorphic nature of CRISPR in phylogenetically related *E. coli* strains makes them potential markers for bacterial typing and evolutionary studies. In the current work, we investigated the occurrence and diversity of CRISPR/Cas system and explored its potential for genotyping. Type I–F CRISPR/Cas systems were found in 413 of 1190 strains of *E. coli* and exhibited the clustering within certain CCs and STs. And CRISPR spacer contents correlated well with MLST types. The divergence analysis of CRISPR showed stronger discriminatory power than MLST, and CRISPR polymorphism was instrumental for differentiating highly closely related strains. The timeline of spacer acquisition and deletion provided important information for inferring the evolution model between distinct serotypes. Identical spacer sequences were shared by strains with the same H-antigen type but not strains with the same O-antigen type. The homology between spacers and antibiotic-resistant plasmids demonstrated the role of Type I–F system in limiting the acquisition of antimicrobial resistance. Collectively, our data presents the dynamic nature of Type I–F CRISPR in *E. coli* of phylogenetic group B2 and provides new insights into the application of CRISPR-based typing in the species.

1. Introduction

Escherichia coli is not only an important commensal usually seen in the gut of human and other warm-blood animals, but also a major pathogen responsible for one million deaths one year owing to intestinal and extraintestinal infections (Leimbach et al., 2013; Wirth et al., 2006). The population structure of the bacterium is highly clonal and is subdivided into seven different phylogenetic groups (A, B1, B2, C, D, E and F) (Clermont et al., 2013). These phylogroups appear to display differences in terms of their metabolic abilities, ecological niches and, especially their virulence or antimicrobial resistance profiles (Bingen et al., 1998). Among them, the phylogenetic group B2 has been significantly associated with extraintestinal infections in human, such as urinary tract infection, newborn meningitis, sepsis and abdominal suppuration (Ewers et al., 2007; Jakobsen et al., 2010). Moreover, multiple investigations have demonstrated that phylogroup B2 is also responsible for a very high proportion of multi-drug resistant *E. coli* (Weissman et al., 2016; Yahiaoui et al., 2015). Mounting evidence has shown that *E. coli* infections related to phylogroup B2 are on the rise, posing a great threat to health (Petersen et al., 2015; Ranjan et al., 2017). To track and control outbreaks caused by *E. coli* belonging to phylogroup B2 and to identify emerging virulent or multi-drug resistant clones, an accurate and efficient typing method is needed.

Over the past few decades, the serotyping typing scheme based on immune mechanism has been used for genotyping pathogenic bacteria, and the method is based on the identification of specific somatic (O) and flagellar (H) antigens (Nuesch-Inderbinen et al., 2018). Although the typing strategy is widely used, it is time-consuming and does not provide a sufficient discriminatory power for outbreak investigations. Furthermore, the method could be not used to deduce phylogenetic relationship between strains. Recently, typing strategies based on DNA sequence are on the increase, including multi-locus sequence typing (MLST), pulse-field gel electrophoresis (PFGE), and whole genome sequencing (WGS). The PFGE typing is performed through analyzing the band patterns of the whole bacterial genome digested with one specific restriction enzyme, which offers an excellent discriminatory power for

* Corresponding author at: Department of Epidemiology, College of Public Health, Zhengzhou University, No. 100 Kexue Avenue, Zhengzhou, Henan Province 450001, China.
E-mail addresses: sychen@zzu.edu.cn (S. Chen), gcduan@zzu.edu.cn (G. Duan).

https://doi.org/10.1016/j.meegid.2019.103916
Received 11 March 2019; Received in revised form 22 May 2019; Accepted 7 June 2019
Available online 10 June 2019

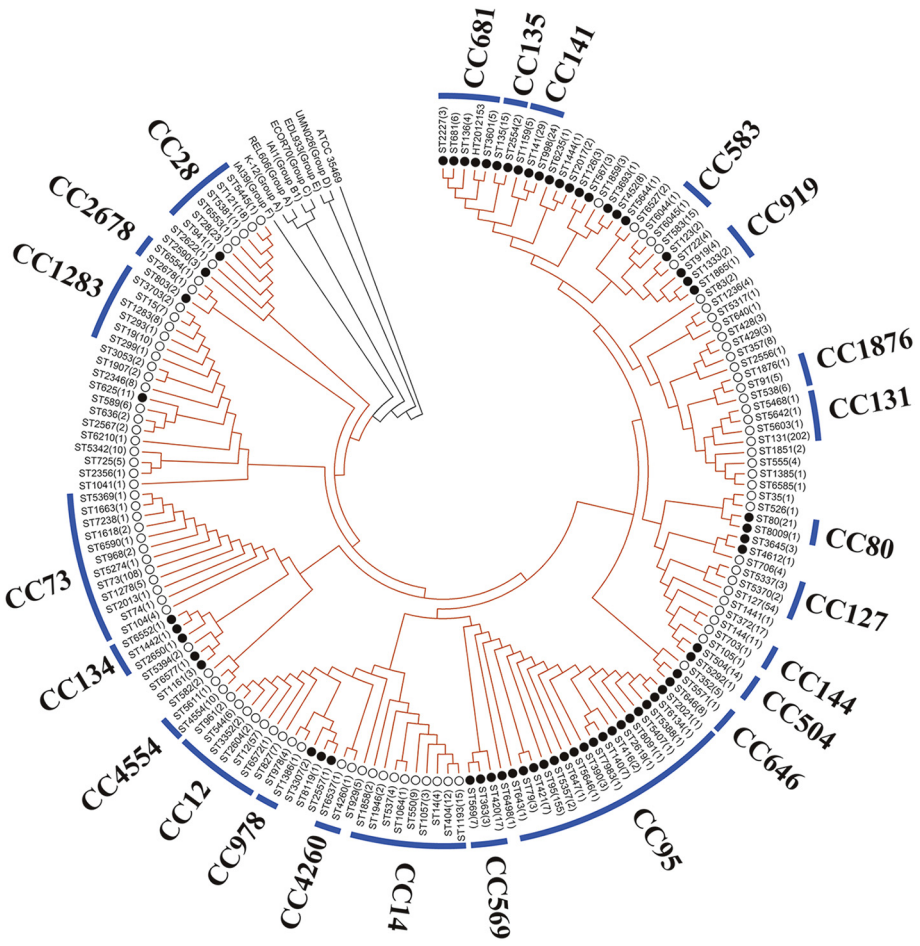

Fig. 1. Distribution of CRISPR/Cas system across MLST in *E. coli* of phylogenetic group B2. The solid and hollow indicate the STs with Type I-F system and without them, respectively. The red line represents strains within the phylogenetic group B2.

lineages analysis (Poulsen et al., 2017). Nevertheless, the typing method depends on specialized instrument and analysis software and could not be exchanged between laboratories worldwide. Although the WGS typing scheme is considered as the "gold" standard for strain typing, the approach is still not applicable to most laboratories due to the lack of expediency and convenience required for routine trace (Lee et al., 2017). MLST is the most widely used typing technique and the typing result is exchangeable, but it requires the sequencing of multiple house-keeping genes and is expensive (Bousquet et al., 2018).

The clustered regularly interspaced short palindromic repeats (CRISPR) coupled with associated *cas* genes, constitute an adaptive immunity system against invasion by mobile genetic elements (MGEs), including viruses and plasmids (Barrangou et al., 2007; Brouns et al., 2008). The system has been identified in most archaea and almost half of bacteria (Grissa et al., 2007). The CRISPR array is composed of a series of (25- to 47-bp) direct repeats separated by variable DNA sequences of similar size, called as spacers, which are derived from foreign DNA (Kunin et al., 2007). The basic ability of CRISPR immunity is to acquire a novel spacer from foreign DNA in a polarized manner at the proximal end of leader sequence (Alkhnbashi et al., 2016). Therefore, positional information of spacers in the CRISPR array reflects a chronological order of the exposure to invasive DNA. Spacer at leader-proximal end was recently acquired and represented recent MGE exposures. Likewise, diversified CRISPR between strains may arise due to selective pressures of diversified MGEs. Thus, the polymorphic nature of CRISPR allows for the application of CRISPR as a high-resolution biomarker for genotyping. At present, CRISPR-based typing methods have been employed for many bacterial species, such as *Yersinia* and *Salmonella* (Barros et al., 2014; Li et al., 2018). In *Escherichia coli,* two subtypes of CRISPR/Cas systems have been deciphered, Type I-E and Type I–F (Diez-Villasenor et al., 2010). Several investigations have reported that Type I–F system is only found in phylogroup B2 and is constitutively expressed (Almendros et al., 2012; Touchon et al., 2011). However, investigations in-depth analyzing the diversity of Type I–F CRISPR/Cas system in *E. coli* of the phylogenetic group B2 are still limited (Aydin et al., 2017).

In this study, we analyzed the occurrence and diversity of Type I–F CRISPR/Cas system in phylogroup B2 *E. coli* to characterize the CRISPR components, exploring the divergence of CRISPR between strains, and evaluated the potential of CRISPR for genotyping *E. coli*.

2. Materials and methods

2.1. Sequence collection

As of May 5, 2018, complete and draft *E. coli* genomes belonging to phylogroup B2 were downloaded from NCBI database. Quality control of genome assemblies was done based on three criteria: (i) for those strains with more than one uploaded genome sequences, only the genomes with the least scaffolds were included, (ii) those genomes originating from single-end short reads were removed as these genomes were likely to lack or misassemble CRISPR arrays owing to the abundance of repeats, and (iii) those genomes where CRISPR arrays and *cas* genes were interrupted owing to contig gaps were also excluded. The NCBI accession numbers of all genomes analyzed in the current work were listed in Table S1.

2.2. Multi-locus sequence typing

MLST types of available *E. coli* genomes were determined *in silico* using the Achtman's MLST scheme (Wirth et al., 2006). The typing scheme utilized conserved sequences within seven pairs of housekeeping genes (*adk*, *icd*, *fumC*, *purA*, *gyrB*, *icd*, *recA*), and the results of MLST for each genome were acquired through the CGE web server (Larsen et al., 2012). Each MLST type was assigned to one clonal complexes (CC) or one singleton using eBURST V3, as previously described (Yu et al., 2018). In order to assure that B2 genomes analyzed did not include genomes misidentified as B2, sequences of seven conserved genes mentioned above were concatenated to construct a phylogenetic tree using neighbor-joining method (Fig. 1). Eight strains that have been previously defined as non-B2 phylogroup were used as outgroups (REL606 and K-12 for phylogroup A, IAI1 for phylogroup B1, UMN026 for phylogroup D, ECOR70 for phylogroup C, EDL933 for phylogroup E, IAI39 for phylogroup F and ATCC 35469 for *Escherichia fergusonii*) (Clermont et al., 2013). The construction of neighbor-joining tree was performed by MEGA 7.0 version.

2.3. Identification of CRISPR/Cas system

CRISPR arrays were identified by CRISPR Recognition Tool with the following setting: repeat length 27–30 nt, spacer length 28–35 nt, minimum repeats per array 2, and search window 8 nt. Irregular repeats was manually proofed if necessary. In parallel, Type I–F *cas* genes were identified through BLASTn search option using phylogroup B2 *E. coli* strain PSUO2 *cas* genes as reference sequences and a minimum of 85% DNA sequence identity. Type I–F *cas* gene clusters were annotated in strain PSUO2 (*cas1, cas2–3, csy1, csy2, csy3, and csy4* encoded as locus tag: ACJ74_19735, ACJ74_19730, ACJ74_19725, ACJ74_19720, ACJ74_19715, ACJ74_19710).

2.4. Subtyping and visualization of CRISPR arrays across each genome

The CRISPRtionary tool was applied to identify “CRISPR types”. The web tool could compare CRISPR alleles and assign each spacer one unique number based on sequence identity. The type number was then acquired for each unique spacer combination within a CRISPR locus with the help of in-house R script. The numbering systems of CRISPR3 and CRISPR4 locus were different. Each unique combination of CRISPR3 and CRISPR4 type number was given a combined CRISPR type (CCT). The visualization of CRISPR arrays was accomplished by the CRISPRstudio tool (Dion et al., 2018).

2.5. Serotype and plasmid replicon typing

The serotypes for strains harboring Type I–F system were predicted through the SeroTypeFinder web tool (Identity ≥80%, coverage rate ≥60%) (Joensen et al., 2015). The types of plasmid replicon were identified through the PlasmidFinder web tool (Identity ≥95%, coverage rate ≥60%) (Carattoli et al., 2014).

2.6. Identification of CRISPR Protospacer

The putative origin of B2 *E. coli* CRISPR spacers was analyzed by recognition of protospacers in available genome sequences in NCBI databases. BLASTn and CRISPRtarget were utilized for homology searching for potential protospacers (Biswas et al., 2013). A “strong” protospacer match was considered when two compared sequence showed > 84% identity (minimum of 27/32 matching nucleotides).

3. Results

3.1. Occurrence and characterization of Type I–F CRISPR/Cas system in E. coli of phylogenetic group B2

A total of 1190 non-redundant *E. coli* genomes belonging to phylogroup B2, including 68 complete and 1122 draft genomes, were collected from NCBI database based on the inclusion and exclusion criteria. According to the Achtman's MLST scheme, all genomes were identified to belong to 161 MLST types (STs), covering 24 ST clonal complexes and 56 singletons. Strain HT2012153 could not be assigned a ST, as the strain produced a novel allele combination yet to be assigned a ST number. Strain HT2012153 was assigned to CC681 through eBURST analysis.

All strains were further analyzed for the occurrence of CRISPR/Cas system through *in silico* analysis. Of them, Type I–F system was identified in 413 strains (34.71%, 413/1190), whereas no Type I–E system was detected. Notably, the Type I–F system-positive isolates exhibited the clustering within particular STs (Fig. 1). Among 24 clonal complexes, only CC95, CC141, CC80, CC134, CC135, CC504, CC569, CC646, CC681, CC919, and ST941 of CC28 were Type I–F system-positive. Regarding 56 singletons, only 20 singletons, including ST123, ST126, ST625, ST803, ST1442, ST1444, ST1161, ST2017, ST2557, ST2590, ST3307, ST3645, ST3693, ST452, ST4612, ST5644, ST6235, ST6498, ST6577, and ST8119 were Type I–F-positive.

Type I–F CRISPR/Cas system in *E. coli* of phylogenetic group B2 was identified to be composed of two CRISPR loci (CRISPR3 and CRISPR4) and six *cas* genes (*cas1*, *cas2–3, csy1, csy2, csy3* and *csy4*) (Fig. S1). For Type I–F system-negative isolates, orphan CRISPR3–4 array was observed in corresponding regions. The lengths of CRISPR3 and CRISPR4 arrays were both highly variable, ranging from 2 to 38 spacers and from 0 to 34 spacers, respectively, whereas the length of orphan CRISPR3–4 array was conserved, with most having one spacer (Table 1). Comparative analysis of spacer contents revealed 245 different spacers in CRISPR3 array and 190 different spacers in CRISPR4 (Table 1). The average length of all spacers identified in CRISPR3 and CRISPR4 arrays was 32 bp, ranging from 31 to 34 bp. In CRISPR3 arrays, only 53 of 245 different spacers was unique (frequency only once), and in CRISPR4 array, only 37 spacers were unique. These unique spacers were more likely to be found in strains of less common MLST types, such as ST452 and ST1442 (Fig. S3). Although many strains were found to share spacers in the same CRISPR loci, there were no shared spacers found between CRISPR3 and CRISPR4 arrays. Orphan CRISPR array seemed to be less polymorphic than CRISPR3 and CRISPR4, and only 5 different spacers were identified in orphan CRISPR array.

When analyzing the similarities between spacers and mobile genetic elements, we observed that 137 of 435 different spacers (31.49%, 137/435) in CRISPR3 and CRISPR4 array showed significant similarities to foreign DNA (Table S2 and S3). Of them, 101 spacers (73.72%, 101/137) were identified to match plasmids, 36 spacers (26.28%, 36/137) could match phages or prophages. These spacers with matches were observed in 384 strains with Type I–F system. Further plasmid typing analysis showed that 45 of 101 spacers that were homologous to plasmids could match conserved regions of IncFII, IncI1, and IncFIB-type antibiotic-resistant plasmids.

In addition, Type I–F *cas* genes were also screened for *E. coli* strains belonging to phylogroup B2. Among Type I–F system-positive isolates, 7 strains (three strains for ST95, three strains for ST140 and one strain for ST421) harbored incomplete *cas* gene cluster. As shown in Fig. S2, the insertion of IS (insertion sequences) or transposases broke the integrity of *cas* gene cluster.

3.2. CRISPR spacer diversity across MLST for genotyping highly related strains

Most CRISPR spacer-based dissimilarities among MLST were found

Table 1
General characteristics of CRISPR arrays from *E. coli* of phylogenetic group B2 (n = 1190).

Characteristics	CRISPR locus		
	CRISPR3	CRISPR4	CRISPR3-4
No. of isolates with array	413	413	777
No. of spacers in array			
Range	2–38	0–34	0–2
Avg	8	8	1
Mode	8	6	1
Total of No. spacers	3478	3226	843
No. of different spacers	245	190	5
No. of unique spacers	53	37	0
Spacer length (bp)			
Avg	32	32	32
Min	32	31	32
Max	34	33	32
No. of different array	110	128	5
No. of unique array	71	95	0

towards the ancestral end, and shared ancestral end implied that they may originate from one common ancestor, reflecting their evolutionary relatedness. Based on the principle that all strains within one given genotype shared at least three most ancestral spacers in CRISPR3 and CRISPR4 array, respectively, 413 strains with Type I–F system were divided into 23 genotypes (Genotypes A-W). Notably, the typing scheme based on conserved ancestral end largely correlated with ST (Fig. 2). For instance, all strains assigned to genotype A formed a large cluster of 189 isolates, which were all members of CC95 (17 STs) (Fig. 3). Exceptionally, the five STs (ST126, ST1161, ST3307, ST6235, and ST6577) shared conserved ancestral end, but they were not grouped together based on MLST. Another exception was that ST363 strains were assigned to the same genotype as the ST452 and ST3693, although they exhibited relatively distant phylogenetic distance based on MLST.

Acquired spacers later may differ due to the differential exposure to mobile genetic elements, indirectly reflecting the divergence between strains evolution. According to the divergence of spacer contents and organizations, 110 CCT3 (CRISPR3 pattern), 128 CCT4 (CRISPR4 patterns), and 182 CCT (combined CRISPR patterns) were identified within all strains belonging to phylogroup B2. In order to compare the potential of CRISPR and MLST for genotyping highly related strains, we explored the CC141 group, including ST141, ST998, and ST1159, which were grouped together based on MLST. We identified 31 distinct CCTs among 53 genomes within the group, although they were assigned to the same genotype C (Fig. 4). The deletion of 8 spacers in CRISPR4 array (at position 14–21) was sufficient to differentiate ST1159 and the two STs (ST141 and ST998). ST141 strains showed highly similar CRISPR spacer compositions and arrangements to ST998, but there were no shared CCTs found between the two STs. Similarly, the resolution of CRISPR-based genotyping was also explored in the CC681 group, including ST136, ST681, ST2227, ST3601, and strain HT2012153, which were assigned to the same genotype D. From the analysis, 16 CCTs were found among 19 strains within the group (Fig. 5). ST3601 and strain HT2012153 differed from other STs by the acquisition of two spacers at position 24–25 in CRISPR3 array. Three ST2227 strains displayed completely identical CRISPR3 pattern, which could differentiate ST2227 strains from ST136 and ST681. ST136 and ST681 strains seemed to show a closer relationship, and the acquisition or deletion of specific spacers was not enough to differentiate the two STs. However, both ST136 and ST681 contained unique CRISPR types and could be used to differentiate each other. Additionally, the polymorphism of Type I–F CRISPR spacer were also instrumental for differentiating other closely related strains, which showed greater discriminatory power than MLST (Fig. S3).

3.3. Evolutionary divergence of CRISPR spacer among serotypes

When analyzing the divergence of CRISPR between different serotypes, we noted that spacer arrangement and organization seemed to correlate well with isolates sharing the same H-antigen type but not necessarily with isolates containing the same O type. To further test this finding, CRISPR spacer contents of 1 O109:H5, 4 O2:H5, 3 O2:H1, 5 O2:H6, and 2 O2:H7, 3 O83:H1, and 3 O83:H33 strains were analyzed. As shown in Fig. 6, O2:H1 did not share any recently acquired spacer in CRISPR3 with O2:H5, O2:H6, and O2:H7, although they had common ancestral spacers in CRISPR3 and CRISPR4. However, O2:H1 clustered together with O83:H1 and shared almost completely identical spacers in CRISPR3 and CRISPR4. Furthermore, O83:H1 shared only ancestral spacers in CRISPR3 and CRISPR4 with O83:H33, but recently acquired spacers were completely different. On the other hand, O2:H5 shared recently acquired spacers in CRISPR3 with O109:H5 but not with O2:H7. Again, subtype I–F *cas* system genes from 21 strains mentioned above were concatenated to constructed one neighboring phylogenetic tree, which was consistent with the evolutionary trend of CRISPR spacer (Fig. 7).

Recently acquired spacers were always incorporated into the proximal end of leader sequence, but the deletion of spacers often, not always, occurred in ancestral end (Makarova et al., 2006). Based on the acquisition and deletion of specific spacers, the time order of the evolution between serotypes could be inferred. All strains among ST95 clonal complex could be further divided into two sub-genotypes (Genotype A1 and A2) based on one putative spacer acquisition at position 11 in CRISPR3 array (Fig. 3). Serotype results indicated that genotype A2 corresponded to the two serotypes retaining the same H5-antigen type (O2:H5 and O109:H5) whereas genotype A1 corresponded to the seven serotypes (O1:H7, O2:H7, O18:H7, O45:H7, NT:H7, O25:H4 and O2:H4). It may be predicted that one spacer acquisition occurred during the evolution of strain serotypes, resulting in those strains that assigned to genotype A2. In parallel, all ST135 and ST2554 strains were assigned to the same genotype B, which contained the same H1-antigen (Fig. 8a). These strains could be further classified into three sub-genotypes (genotype B1-B3) based on the acquisition and deletion of spacer. Genotype B2 and B3 differed from B1 by one spacer acquisition at position 34 in CRISPR4. And genotype B2 differed from B3 by the deletion of four spacers at position 7–10 in CRISPR4. Serotype results showed that genotype B2 and B3 both corresponded to serotype O2:H1 whereas genotype B1 corresponded to serotype O83:H1. Therefore, one putative evolution model from serotype O2:H1 to O83:H1 could be inferred (Fig. 8b). Serotype O2:H1 strains assigned to ST135 (Genotype B1) was the predecessor, followed by O83:H1 strains assigned to ST135 (Genotype B2), and O83:H1 strains assigned to ST2554 (Genotype B3).

4. Discussion

The characterization of CRISPR/Cas system in *E. coli* has been well described, and four CRISPR loci (CRISPR1 and CRISPR2 for Type I-E, and CRISPR3 and CRISRP4 for Type I–F) have been identified in previous reports (Diez-Villasenor et al., 2010; Touchon et al., 2012; Touchon and Rocha, 2010). One study has demonstrated that Type I–F CRISPR/Cas system is only present in the phylogenetic group B2, but Type I-E system is widely distributed in the remaining phylogenetic groups of the specie (Touchon et al., 2011). In the past two decades, the adaptive immunity mechanism underlying the Type I-E system has been well clarified, and its CRISPR spacer polymorphisms have also been applied for the detection of shiga-toxin producing *E. coli* strains with specific serotypes (Delannoy et al., 2012a,b). As for Type I–F CRISPR/Cas system, its divergence in CRISPR array is regard as useful for reconstituting the evolutionary history of *Yersinia* species and genotyping the causative agents (Barros et al., 2014; Koskela et al., 2015; Vergnaud et al., 2007). Other research groups have also explored its value for typing other bacteria species, such as *Helicobacter cinaedi* and

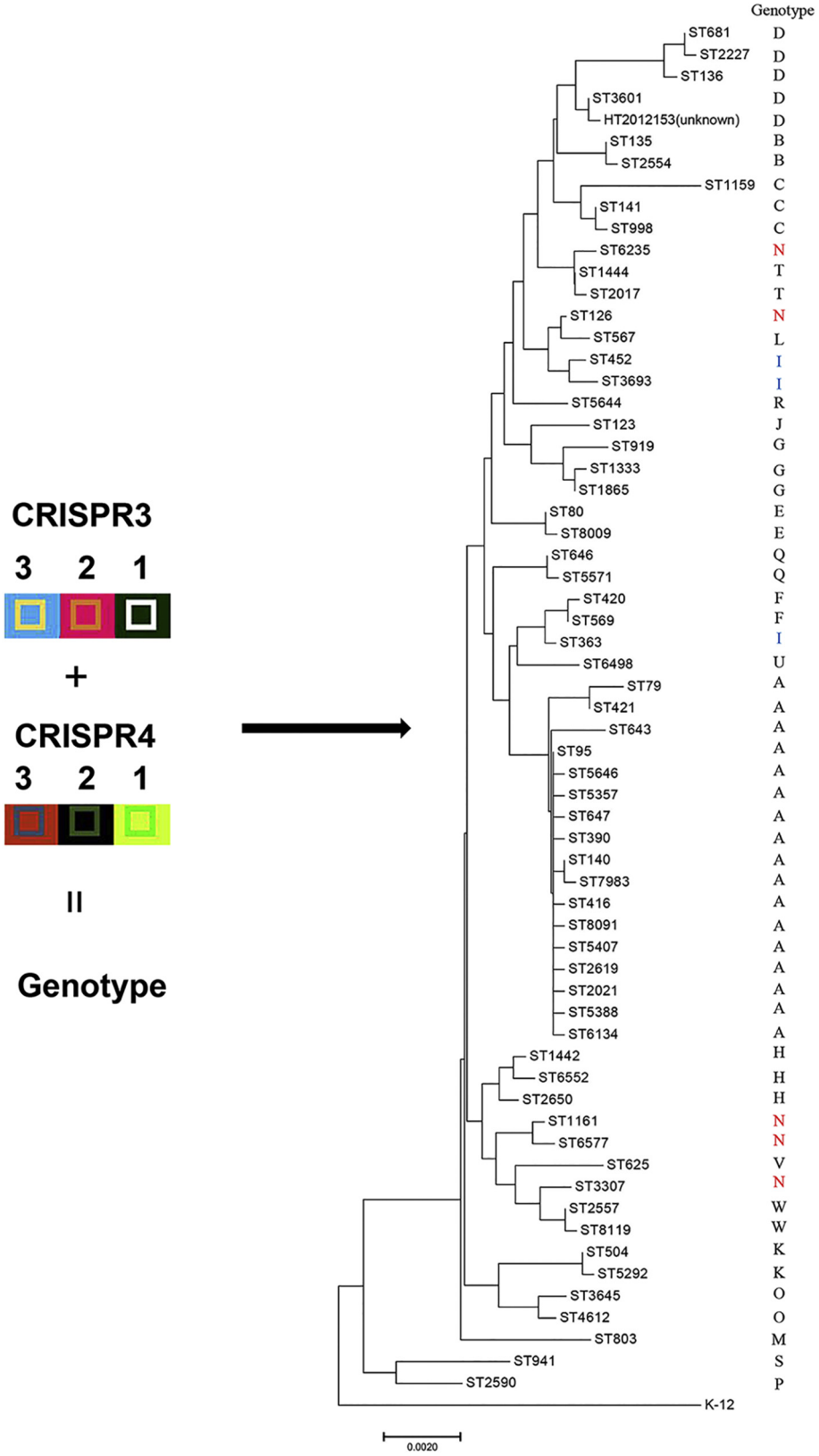

Fig. 2. Association between clustering by shared ancestral end and MLST. The sequences of seven house-keeping genes (*adk, icd, fumC, purA, gyrB, icd, recA*) from strains representing 62 STs and one unassigned ST are concentrated to construct a neighbor-joining tree. The *Escherichia coli* K-12 is used as one outgroup. The CRISPR genotype per ST is marked in the corresponding position.

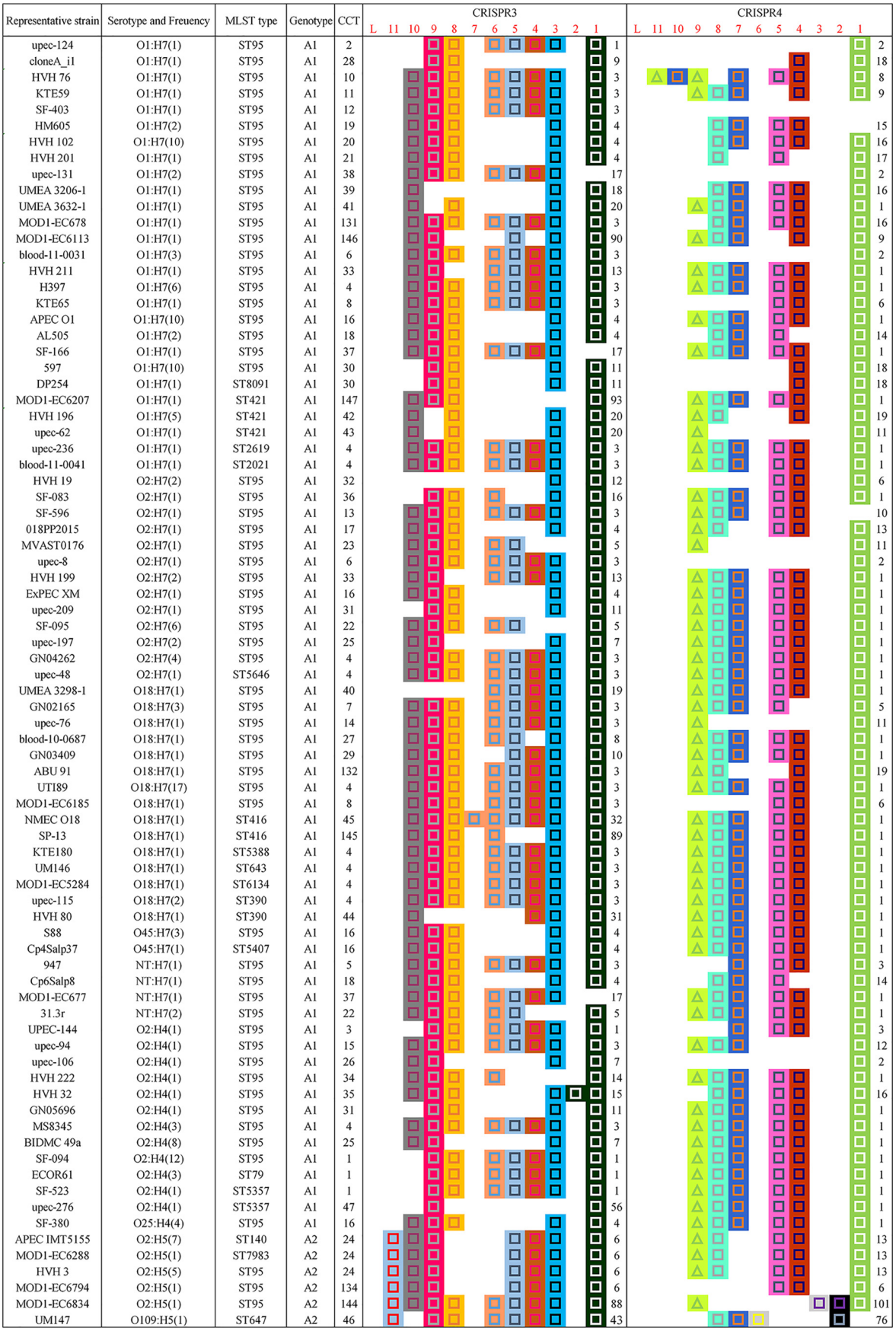

Fig. 3. CRISPR spacer arrangement and polymorphisms across ST95 clonal complex. Spacers are shown as squares uniquely colored by spacer sequence, with different icons representing spacer length. Spacer numbering is initiated at the ancestral end (right) towards the most recently acquired spacers per strain towards the left. The left side represents CRISPR3 and the right represents CRISPR4 for same strain in the same line. Column L symbolizes the leader position. CRISPR type (CCT) assignment is displayed in the CCT column, and CRISPR3 and CRISPR4 allele numbers are exhibited in their corresponding column.

Representative strain	MLST type and Freuency	Genotype	CCT	CRISPR3	CRISPR4
				L 9 8 7 6 5 4 3 2 1	L 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1
UPEC-70	ST1159(1)	C	62	27	59
UPEC-89	ST1159(1)	C	63	27	60
MOD1-EC5468	ST1159(1)	C	64	50	82
MOD1-EC6843	ST1159(1)	C	143	85	93
Blood-09-0829	ST141(7)	C	65	27	35
H223	ST141(1)	C	66	27	36
HM65	ST141(1)	C	67	27	37
HVH 38	ST141(1)	C	68	27	38
KTE145	ST141(1)	C	69	27	39
KTE15	ST141(1)	C	70	27	53
KTE16	ST141(1)	C	71	27	40
KTE169	ST141(1)	C	73	27	42
KTE183	ST141(1)	C	74	27	43
KTE215	ST141(1)	C	76	27	45
KTE224	ST141(1)	C	77	27	46
KTE43	ST141(1)	C	78	27	47
KTE72	ST141(1)	C	79	27	48
KTE86	ST141(1)	C	80	27	49
KTE93	ST141(1)	C	81	27	50
M17-1	ST141(1)	C	82	27	51
PSUO2	ST141(1)	C	83	27	52
SCB11	ST141(1)	C	84	28	61
Chicken Fecal	ST141(1)	C	85	28	62
F356	ST141(1)	C	141	49	114
9.1649	ST998(1)	C	87	27	54
Fec 67	ST998(1)	C	88	27	55
MOD1-EC5941	ST998(12)	C	89	27	56
UCD-JA19	ST998(1)	C	91	27	58
1408270010	ST998(4)	C	92	49	56
MOD1-EC711	ST998(1)	C	140	27	115
PairA_ch_1	ST998(3)	C	148	110	125

Fig. 4. CRISPR spacer organizations and arrangements in the group of ST141, ST998 and ST1159.

Representative strain	MLST type and Freuency	Genotype	CCT	CRISPR3	CRISPR4
				L 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1	L 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1
HT2012163	ST136(1)	D	176	98	97
KD1	ST136(1)	D	123	26	34
MOD1-EC6070	ST136(1)	D	75	107	116
MOD1-EC890	ST136(1)	D	180	102	109
H001	ST681(1)	D	120	45	78
ECOR65	ST681(1)	D	121	68	90
MOD1-EC6151	ST681(1)	D	173	95	122
HT2012099	ST681(1)	D	133	45	111
upec-116	ST681(1)	D	122	46	79
50870281	ST681(1)	D	119	44	77
MOD1-EC5824	ST2227(1)	D	172	94	119
MOD1-EC5884	ST2227(2)	D	171	94	112
HT2012153	ST3601(1), unknown(1)	D	86	106	113
HT2012055	ST3601(2)	D	182	104	113
HT2012DB03	ST3601(1)	D	90	105	113
HT2012088	ST3601(1)	D	72	108	117

Fig. 5. CRISPR spacer organizations and arrangements in the group of ST136, ST681, ST2227, ST3601 and one unassigned ST.

Representative strain	O-type	H-type	MLST type	Genotype	CCT	CRISPR3	CRISPR4
						L 13 12 11 10 9 8 7 6 5 4 3 2 1	L 34 33 32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1
upec-48	O2	H7	ST5646	A1	4	3	1
GN04262	O2	H7	ST95	A1	4	3	1
UM147	O109	H5	ST647	A2	46	43	76
MOD1-EC6834	O2	H5	ST95	A2	144	88	101
MOD1-EC6794	O2	H5	ST95	A2	134	6	6
HVH 3	O2	H5	ST95	A2	24	6	13
MOD1-EC6288	O2	H5	ST7983	A2	24	6	13
MOD1-EC5941	O2	H6	ST998	C	89	27	56
MOD1-EC5468	O2	H6	ST1159	C	64	50	82
MOD1-EC6843	O2	H6	ST1159	C	143	85	93
PSUO2	O2	H6	ST141	C	83	27	52
SCB11	O2	H6	ST141	C	84	28	61
APEC O2	O2	H1	ST135	B1	52	24	24
MOD1-EC6529	O2	H1	ST135	B1	135	24	44
MOD1-EC6891	O2	H1	ST135	B1	136	24	41
LF82	O83	H1	ST135	B2	53	24	25
NRG 857C	O83	H1	ST135	B2	54	24	26
MOD1-EC5478	O83	H1	ST135	B2	58	24	30
KTE126	O83	H33	ST567	L	94	39	72
KTE47	O83	H33	ST567	L	95	40	73
MOD1-EC5680	O83	H33	ST567(1)	L	96	40	74

Fig. 6. CRISPR spacer organizations and arrangements from 1 O109:H5, 4 O2:H5, 3 O2:H1, 5 O2:H6, and 2 O2:H7, 3 O83:H1, and 3 O83:H33 strains.

Pseudomonas aeruginosa (Tomida et al., 2017; van Belkum et al., 2015). However, little analysis has been published for describing the characterization of Type I–F CRISPR/Cas system with detail and exploring its potential for typing strains in *E. coli* strains of phylogenetic group B2.

In the current work, we determined the remarkable polymorphism of CRISPR spacer among Type I–F system-positive strains, which could be sufficient for differentiating highly related strains. In most cases, multiple CCTs were found in one CC or single ST, which allows further discrimination between strains of certain CCs or STs. One study in *Clostridium difficile* also demonstrated that CRISPR sequences harbored greater variability than MLST, which is useful for identifying *Clostridium difficile* strains harboring specific virulent markers (Andersen et al., 2016). Moreover, we found that CRISPR spacer contents correlated with the MLST types. Most of the spacer sequences identified among certain CCs or STs, aside from ancestral spacer, were not found among other CCs or STs, which suggested their STs

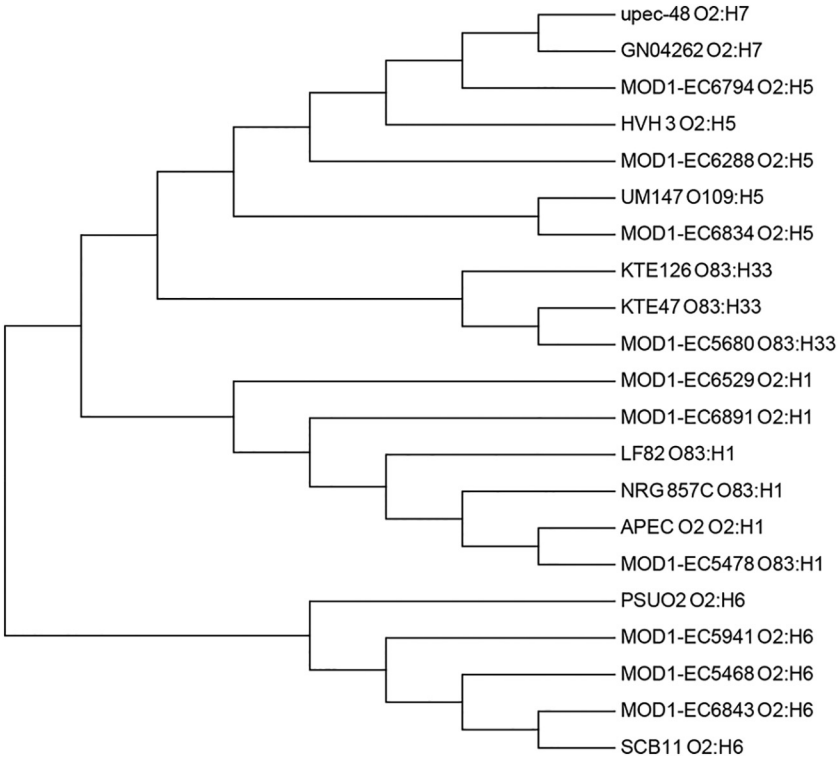

Fig. 7. Phylogenetic tree of subtype I-F cas genes from 1 O109:H5, 4 O2:H5, 3 O2:H1, 5 O2:H6, and 2 O2:H7, 3 O83:H1, and 3 O83:H33 strains.

association. Another study focusing on Type I-E CRISPR polymorphisms in *E. coli* have also showed that highly related strains of MLST phylogenetic distance < 0.005 often exhibited similar or identical CRISPR pattern whereas more distantly related strains of MLST phylogenetic distance > 0.01 displayed completely different CRISPR patterns (Touchon et al., 2011). Therefore, CRISPR polymorphisms analysis seemed to be applicable for further analysis of strains within CCs or STs to better understand the epidemiological relationships of STs in strains having temporal and spatial connection. However, it was notably that CRISPR may be acquired even prior to MGEs exposure through horizontal transfer. In the occurrent work, the unconformity between ancestral-end-based clustering and the MLST analysis supported the speculation. Thus, CRISPR-based analysis should be used with caution to reconstruct the phylogenetic relationship between isolates of distantly related STs. Additionally, the absence of CRISPR in some strains of interest may preclude the widespread exploration of this typing technology. The combination of multiple typing technologies may be an alternative in practical applications.

H-antigen alleles seem to be more phylogenetically stable whereas O-antigen loci are shuffled in the evolution of some *E. coli* lineages. Our data showed that strains of different O groups combined with the identical H antigen exhibited almost completely identical arrays (O2:H1 and O83:H1, O2:H5 and O109:H5), while arrays of isolates of the same O type with different H alleles appeared to be unrelated (O2:H1, O2:H5, O2:H6 and O2:H7). Other investigations also demonstrated the association between CRISPR spacer components and H-antigen alleles. Delannoy et al. developed one CRISPR-based real-time PCR assays to specially detect the presence of isolates with O26:H11,

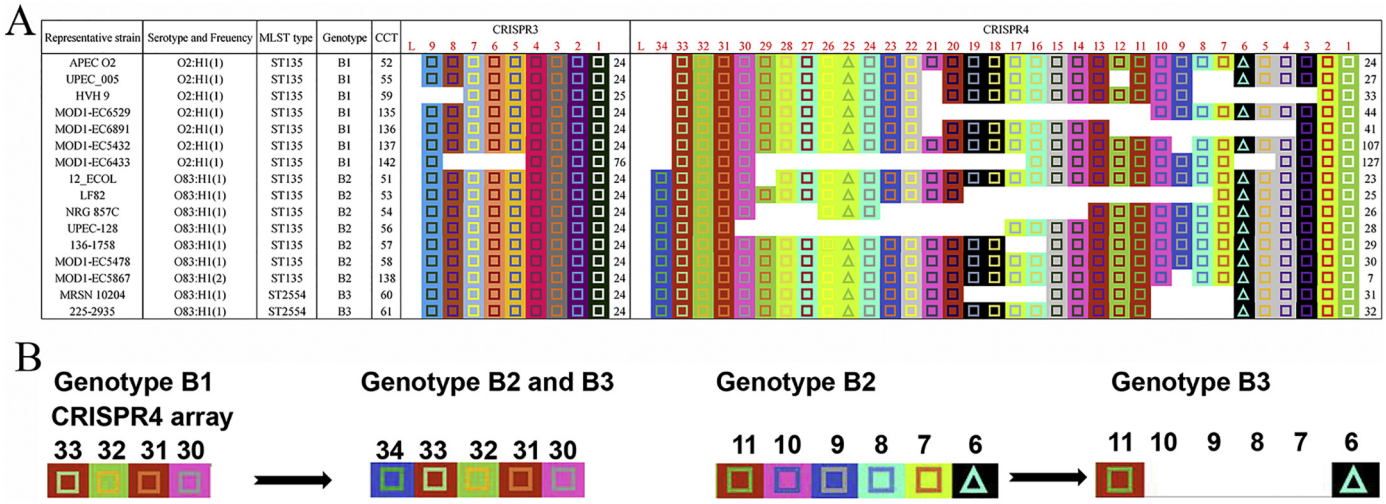

Fig. 8. CRISPR array in the group of ST135 and ST2554. CRISPR spacer organizations and arrangements in the group of ST135 and ST2554 (A). Three CRISPR subgenotypes in the group of ST135 and ST2554 (B).

O45:H2, O103:H2, O104:H4, O145:H28 and O157:H7 (Delannoy et al., 2012b). Nevertheless, they found several cross-reaction within isolates with the same H-antigen type, such as the cross-reaction between O145:H28 and O28:H28, and the cross-reaction among O103:H2, O45:H2, O128:H2 and O145:H2. Moreover, Yin et al. also demonstrated the conservation of CRISPR contents within strains of the identical H type, including H7, H2 and H11 (Yin et al., 2013). In addition, strains between serotype O2:H1 and O83:H1 are indistinguishable based on Type I–F *cas* genes, and so are strains between O2:H5 and O109:H5, which further highlighting the linkage between CRISPR contents and H type. CRISPR evolve by polarized acquisition of novel spacer in the proximal end of leader sequence, which reflects a chronological record of MGEs exposure and provides one additional tool for evolutionary study. The directionality of novel spacer acquisition has allowed the construction of a hypothetic evolutionary model of *Yersinia pestis*, which could be used for inferring the transmission route of *Y. pestis* strains (Barros et al., 2014). Similarly, putative spacer acquisition events could be applied for reconstituting the evolutionary timeline of different serotypes, which is complementary to the limitations of serotype typing that could not infer phylogenetic relationship.

CRISPR/Cas system is one RNA-mediated immunity system, providing protection against invading MGEs (Garneau et al., 2010). Our results showed that Type I–F system-positive isolates were generally clustered in some distinct clonal complexes and some specific STs. CC131, CC127, CC95, CC73 and CC141 are well known as the dominated clonal complexes responsible for extra-intestinal pathogenetic *E. coli* (ExPEC), of which only CC95 and CC141 are Type I–F system-positive whereas the remaining CCs are Type I–F system-negative. In fact, these Type I–F system-positive CCs have been reported to be often associated with low-level antibiotic-resistance, such as CC95 (Alghoribi et al., 2015), whereas Type I–F system-negative CCs are related to high-level antimicrobial resistance, such as CC131 (Colpan et al., 2013). Type I–F system in *E. coli* has been experimentally demonstrated to be persistently expressed, which provide natural adaption immunity against invasive DNA (Almendros et al., 2012). The basic ability to prevent the incorporation of MGEs into the host enables bacteria to survive in phage-rich environment. However, it may come at a cost of inhibiting the acquisition of other beneficial genes (including virulent and antibiotic-resistant genets) through HGT. Thus, the absence or presence of Type I–F system in bacteria may affect the distribution of virulent or antibiotic-resistance genes. Moreover, one study has revealed that presence of Type I–F CRISPR/Cas system is associated with antimicrobial susceptibility in *Escherichia coli* (Aydin et al., 2017). Additionally, multiple matches between spacers and antibiotic-resistant plasmid were also observed in our work. These data further demonstrated the role of Type I–F system in limiting the acquisition of antibiotic resistance. The integrity of *cas* genes cluster has already been proved to be indispensable for the action of Type I–F CRISPR/Cas system (Vorontsova et al., 2015). The insertion of IS or transposase within *cas* genes cluster could impair the interference mechanism, which seems to be beneficial for the acquisition of MGEs (Chen et al., 2018). In other words, the presence of IS or transposases in Type I–F system within the ST95 clonal complex will help the clonal population acquire these genetic elements with survival advantage. Taken together, when facing strong selective pressure, bacteria may adjust themselves to one state that is conductive to survival, such as regulating the activity of CRISPR/Cas system or lost them.

5. Conclusion

In summary, the above findings extend our understanding for the role of Type I–F CRISPR/Cas system in *E. coli* of phylogenetic group B2. And CRISPR spacer polymorphism displays the potential for typing B2 strains and provides some clues for inferring the phylogenetic relationship between distinct serotypes. H-antigen alleles seem to be more phylogenetically stable than O antigen since CRISPR contents are shared by isolates containing the same H type but not by isolates with the same O group. The widespread presence of spacers matching antibiotic-resistant plasmids hints that Type I–F CRISPR/Cas system may play an important role in hindering the acquisition of antibiotic resistance. Accordingly, Type I–F CRISPR/Cas system is an important factor in B2 *E. coli* evolution, and provide important insights into genome microevolution.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2019.103916.

Acknowledgments

The work was funded by the National Science and Technology Specific Projects (2018ZX10301407, 2013ZX10004607) and Henan Province University Science and Technology Innovation Talent Projects (17HASTIT045).

Author contributions statements

Jinzhao Long and Guangcai Duan designed the study, Jinzhao Long, Yake Xu analyzed data and wrote the paper. Haiyan Yang, Yuanlin Xi and Liuyang Ou collected some data. All authours read and approved the final manuscript.

Declaration of Conflict Interests

The authors declare no conflict of interest.

References

Alghoribi, M.F., Gibreel, T.M., Farnham, G., Al Johani, S.M., Balkhy, H.H., Upton, M., 2015. Antibiotic-resistant ST38, ST131 and ST405 strains are the leading uropathogenic Escherichia coli clones in Riyadh, Saudi Arabia. J. Antimicrob. Chemother. 70, 2757–2762.

Alkhnbashi, O.S., Shah, S.A., Garrett, R.A., Saunders, S.J., Costa, F., Backofen, R., 2016. Characterizing leader sequences of CRISPR loci. Bioinformatics 32, i576–i585.

Almendros, C., Guzman, N.M., Diez-Villasenor, C., Garcia-Martinez, J., Mojica, F.J., 2012. Target motifs affecting natural immunity by a constitutive CRISPR-Cas system in Escherichia coli. PLoS ONE 7, e50797.

Andersen, J.M., Shoup, M., Robinson, C., Britton, R., Olsen, K.E., Barrangou, R., 2016. CRISPR diversity and microevolution in Clostridium difficile. Genome. Biol. Evol. 8, 2841–2855.

Aydin, S., Personne, Y., Newire, E., Laverick, R., Russell, O., Roberts, A.P., Enne, V.I., 2017. Presence of type I-F CRISPR/Cas systems is associated with antimicrobial susceptibility in Escherichia coli. J. Antimicrob. Chemother. 72, 2213–2218.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., Horvath, P., 2007. CRISPR provides acquired resistance against viruses in prokaryotes. Science 315, 1709–1712.

Barros, M.P., Franca, C.T., Lins, R.H., Santos, M.D., Silva, E.J., Oliveira, M.B., Silveira-Filho, V.M., Rezende, A.M., Balbino, V.Q., Leal-Balbino, T.C., 2014. Dynamics of CRISPR loci in microevolutionary process of Yersinia pestis strains. PLoS ONE 9, e108353.

Bingen, E., Picard, B., Brahimi, N., Mathy, S., Desjardins, P., Elion, J., Denamur, E., 1998. Phylogenetic analysis of Escherichia coli strains causing neonatal meningitis suggests horizontal gene transfer from a predominant pool of highly virulent B2 group strains. J. Infect. Dis. 177, 642–650.

Biswas, A., Gagnon, J.N., Brouns, S.J., Fineran, P.C., Brown, C.M., 2013. CRISPRTarget: bioinformatic prediction and analysis of crRNA targets. RNA Biol. 10, 817–827.

Bousquet, A., Bugier, S., Larreche, S., Bigaillon, C., Weber, P., Delacour, H., Valade, E., De Briel, D., Merens, A., 2018. Clinical isolates of Escherichia coli solely resistant to mecillinam: prevalence and epidemiology. Int. J. Antimicrob. Agents 51, 493–497.

Brouns, S.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J., Snijders, A.P., Dickman, M.J., Makarova, K.S., Koonin, E.V., van der Oost, J., 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. Science 321, 960–964.

Carattoli, A., Zankari, E., Garcia-Fernandez, A., Voldby Larsen, M., Lund, O., Villa, L., Moller Aarestrup, F., Hasman, H., 2014. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob. Agents Chemother. 58, 3895–3903.

Chen, S., Liu, H., Liang, W., Hong, L., Zhang, B., Huang, L., Guo, X., Duan, G., 2018. Insert sequences of CRISPR/Cas system regulate horizontal antibiotic gene transfer in Shigella. Int. J. Antimicrob. Agents 53, 109–115.

Clermont, O., Christenson, J.K., Denamur, E., Gordon, D.M., 2013. The Clermont Escherichia coli phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. Environ. Microbiol. Rep. 5, 58–65.

Colpan, A., Johnston, B., Porter, S., Clabots, C., Anway, R., Thao, L., Kuskowski, M.A., Tchesnokova, V., Sokurenko, E.V., Johnson, J.R., Investigators, V., 2013. Escherichia

coli sequence type 131 (ST131) subclone H30 as an emergent multidrug-resistant pathogen among US veterans. Clin. Infect. Dis. 57, 1256–1265.
Delannoy, S., Beutin, L., Burgos, Y., Fach, P., 2012a. Specific detection of enteroaggregative hemorrhagic Escherichia coli O104:H4 strains by use of the CRISPR locus as a target for a diagnostic real-time PCR. J. Clin. Microbiol. 50, 3485–3492.
Delannoy, S., Beutin, L., Fach, P., 2012b. Use of clustered regularly interspaced short palindromic repeat sequence polymorphisms for specific detection of enterohemorrhagic Escherichia coli strains of serotypes O26:H11, O45:H2, O103:H2, O111:H8, O121:H19, O145:H28, and O157:H7 by real-time PCR. J. Clin. Microbiol. 50, 4035–4040.
Diez-Villasenor, C., Almendros, C., Garcia-Martinez, J., Mojica, F.J., 2010. Diversity of CRISPR loci in Escherichia coli. Microbiology 156, 1351–1361.
Dion, M.B., Labrie, S.J., Shah, S.A., Moineau, S., 2018. CRISPRStudio: a user-friendly software for rapid CRISPR Array visualization. Viruses 10, e602.
Ewers, C., Li, G., Wilking, H., Kiessling, S., Alt, K., Antao, E.M., Laturnus, C., Diehl, I., Glodde, S., Homeier, T., Bohnke, U., Steinruck, H., Philipp, H.C., Wieler, L.H., 2007. Avian pathogenic, uropathogenic, and newborn meningitis-causing Escherichia coli: how closely related are they? Int. J. Med. Microbiol. 297, 163–176.
Garneau, J.E., Dupuis, M.E., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadan, A.H., Moineau, S., 2010. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature 468, 67–71.
Grissa, I., Vergnaud, G., Pourcel, C., 2007. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. BMC Bioinfo. 8, 172.
Jakobsen, L., Hammerum, A.M., Frimodt-Moller, N., 2010. Virulence of Escherichia coli B2 isolates from meat and animals in a murine model of ascending urinary tract infection (UTI): evidence that UTI is a zoonosis. J. Clin. Microbiol. 48, 2978–2980.
Joensen, K.G., Tetzschner, A.M., Iguchi, A., Aarestrup, F.M., Scheutz, F., 2015. Rapid and easy in silico serotyping of Escherichia coli isolates by use of whole-genome sequencing data. J. Clin. Microbiol. 53, 2410–2426.
Koskela, K.A., Mattinen, L., Kalin-Manttari, L., Vergnaud, G., Gorge, O., Nikkari, S., Skurnik, M., 2015. Generation of a CRISPR database for Yersinia pseudotuberculosis complex and role of CRISPR-based immunity in conjugation. Environ. Microbiol. 17, 4306–4321.
Kunin, V., Sorek, R., Hugenholtz, P., 2007. Evolutionary conservation of sequence and secondary structures in CRISPR repeats. Genome Biol. 8, R61.
Larsen, M.V., Cosentino, S., Rasmussen, S., Friis, C., Hasman, H., Marvig, R.L., Jelsbak, L., Sicheritz-Ponten, T., Ussery, D.W., Aarestrup, F.M., Lund, O., 2012. Multilocus sequence typing of total-genome-sequenced bacteria. J. Clin. Microbiol. 50, 1355–1361.
Lee, K.I., Morita-Ishihara, T., Iyoda, S., Ogura, Y., Hayashi, T., Sekizuka, T., Kuroda, M., Ohnishi, M., Japan, E.W.G.i., 2017. A geographically widespread outbreak investigation and development of a rapid screening method using whole genome sequences of Enterohemorrhagic Escherichia coli O121. Front. Microbiol. 8, 701.
Leimbach, A., Hacker, J., Dobrindt, U., 2013. E. coli as an all-rounder: the thin line between commensalism and pathogenicity. Curr. Top. Microbiol. Immunol. 358, 3–32.
Li, Q., Wang, X., Yin, K., Hu, Y., Xu, H., Xie, X., Xu, L., Fei, X., Chen, X., Jiao, X., 2018. Genetic analysis and CRISPR typing of Salmonella enterica serovar Enteritidis from different sources revealed potential transmission from poultry and pig to human. Int. J. Food Microbiol. 266, 119–125.
Makarova, K.S., Grishin, N.V., Shabalina, S.A., Wolf, Y.I., Koonin, E.V., 2006. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. Biol. Direct 1, 7.
Nuesch-Inderbinen, M., Morach, M., Cernela, N., Althaus, D., Jost, M., Mausezahl, M., Bloomberg, G., Stephan, R., 2018. Serotypes and virulence profiles of Shiga toxin-producing Escherichia coli strains isolated during 2017 from human infections in Switzerland. Int. J. Med. Microbiol. 308, 933–939.
Petersen, A.M., Halkjaer, S.I., Gluud, L.L., 2015. Intestinal colonization with phylogenetic group B2 Escherichia coli related to inflammatory bowel disease: a systematic review and meta-analysis. Scand. J. Gastroenterol. 50, 1199–1207.
Poulsen, L.L., Thofner, I., Bisgaard, M., Christensen, J.P., Olsen, R.H., Christensen, H., 2017. Longitudinal study of transmission of Escherichia coli from broiler breeders to broilers. Vet. Microbiol. 207, 13–18.
Ranjan, A., Shaik, S., Nandanwar, N., Hussain, A., Tiwari, S.K., Semmler, T., Jadhav, S., Wieler, L.H., Alam, M., Colwell, R.R., Ahmed, N., 2017. Comparative genomics of *Escherichia coli* isolated from skin and soft tissue and other extraintestinal infections. mBio 8, e01070-17.
Tomida, J., Morita, Y., Shibayama, K., Kikuchi, K., Sawa, T., Akaike, T., Kawamura, Y., 2017. Diversity and microevolution of CRISPR loci in helicobacter cinaedi. PLoS ONE 12, e0186241.
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.
Touchon, M., Charpentier, S., Pognard, D., Picard, B., Arlet, G., Rocha, E.P., Denamur, E., Branger, C., 2012. Antibiotic resistance plasmids spread among natural isolates of Escherichia coli in spite of CRISPR elements. Microbiology 158, 2997–3004.
van Belkum, A., Soriaga, L.B., LaFave, M.C., Akella, S., Veyrieras, J.B., Barbu, E.M., Shortridge, D., Blanc, B., Hannum, G., Zambardi, G., Miller, K., Enright, M.C., Mugnier, N., Brami, D., Schicklin, S., Felderman, M., Schwartz, A.S., Richardson, T.H., Peterson, T.C., Hubby, B., Cady, K.C., 2015. Phylogenetic distribution of CRISPR-Cas Systems in Antibiotic-Resistant Pseudomonas aeruginosa. mBio 6, e01796-01715.
Vergnaud, G., Li, Y., Gorge, O., Cui, Y., Song, Y., Zhou, D., Grissa, I., Dentovskaya, S.V., Platonov, M.E., Rakin, A., Balakhonov, S.V., Neubauer, H., Pourcel, C., Anisimov, A.P., Yang, R., 2007. Analysis of the three Yersinia pestis CRISPR loci provides new tools for phylogenetic studies and possibly for the investigation of ancient DNA. Adv. Exp. Med. Biol. 603, 327–338.
Vorontsova, D., Datsenko, K.A., Medvedeva, S., Bondy-Denomy, J., Savitskaya, E.E., Pougach, K., Logacheva, M., Wiedenheft, B., Davidson, A.R., Severinov, K., Semenova, E., 2015. Foreign DNA acquisition by the I-F CRISPR-Cas system requires all components of the interference machinery. Nucleic Acids Res. 43, 10848–10860.
Weissman, S.J., Hansen, N.I., Zaterka-Baxter, K., Higgins, R.D., Stoll, B.J., 2016. Emergence of antibiotic resistance-associated clones among Escherichia coli recovered from newborns with early-onset Sepsis and meningitis in the United States, 2008-2009. J. Pediatric. Infect. Dis. Soc. 5, 269–276.
Wirth, T., Falush, D., Lan, R., Colles, F., Mensa, P., Wieler, L.H., Karch, H., Reeves, P.R., Maiden, M.C., Ochman, H., Achtman, M., 2006. Sex and virulence in Escherichia coli: an evolutionary perspective. Mol. Microbiol. 60, 1136–1151.
Yahiaoui, M., Robin, F., Bakour, R., Hamidi, M., Bonnet, R., Messai, Y., 2015. Antibiotic resistance, virulence, and genetic background of community-acquired Uropathogenic Escherichia coli from Algeria. Microb. Drug Resist. 21, 516–526.
Yin, S., Jensen, M.A., Bai, J., Debroy, C., Barrangou, R., Dudley, E.G., 2013. The evolutionary divergence of Shiga toxin-producing Escherichia coli is reflected in clustered regularly interspaced short palindromic repeat (CRISPR) spacer composition. Appl. Environ. Microbiol. 79, 5710–5720.
Yu, F., Chen, X., Zheng, S., Han, D., Wang, Y., Wang, R., Wang, B., Chen, Y., 2018. Prevalence and genetic diversity of human diarrheagenic Escherichia coli isolates by multilocus sequence typing. Int. J. Infect. Dis. 67, 7–13.