Open Forum Infectious Diseases

MAJOR ARTICLE

Expansion of a Subset Within the C2 Subclade of *Escherichia coli* Sequence Type 131 (ST131) Is Driving the Increasing Rates of Aminoglycoside Resistance

Zoya Hojabri,[1,2] Narges Darabi,[1] Majid Mirmohammadkhani,[3] Hamzeh Rahimi,[4] Romina Hemmati,[2] Zahra Saeedi,[2] Kiarash Roustaee,[2] Pimlapas Leekitcharoenphon,[5] Omid Pajand,[1,2,3,5] and Frank M. Aarestrup[5]

[1]Microbiology Department, Faculty of Medicine, Semnan University of Medical Sciences, Semnan, Iran, [2]Student Research Committee, Faculty of Medicine, Semnan University of Medical Sciences, Semnan, Iran, [3]Social Determinants of Health Research Center, Semnan University of Medical Sciences, Semnan, Iran, [4]Department of Molecular Medicine, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran, and [5]Section for Genomic Epidemiology, National Food Institute, Technical University of Denmark, Lyngby, Denmark

Background. Sequence type 131 (ST131) of *Escherichia coli* is a pandemic clone that drives the increasing rates of antibiotic resistance. While the pervasiveness of ST131 clade C, especially subclades C2 and C1-M27, has been demonstrated in numerous global surveys, no report about the ST131 clades and their virotypes has been published from Iran so far.

Methods. A collection of 73 consecutive ST131 isolates from extraintestinal specimens was investigated for determination of virotypes, antibiotic susceptibility patterns, resistance/virulence determinants, and clade subsets.

Results. Most of the isolates belonged to subclade C2 (33/73; 45.2%), which had the highest virulence factor (VF) scores and resistance rates, followed by C1-M27 (18; 24.6%), C1-non-M27 (14; 19.1%), and A (8; 10.9%). The distinctive profiles of subclade C2 virulence genes were revealed by principle coordinates analysis testing. The distribution of the *hlyA* virulence gene among subclade C2 was not uniform, so that positive strains (21; 63.6%) showed significantly higher rates of resistance ($bla_{\text{CTX-M-15}}$, $bla_{\text{OXA-1}}$, *aac(6')-Ib-cr*, *aac(6')-Ib*, *aac(3)-IIa*) and virulence (*hra*, *tia/hek*, *K5*, *cnf*, *papGII*, *papC*) markers and gentamicin/tobramycin resistance. Virotype C as the most common virotype (34; 46.5%) was predominant among the subclade C1 population, while virotypes E and F (21; 28.7%) were detected among subclade C2, which had the highest VF scores and aminoglycoside resistance rates.

Conclusions. The appearance of virotypes E and F among subclade C2 strains with higher rates of aminoglycoside resistance/virulence gene content shows the shifting dynamics of this pandemic clone in response to antibiotic selection pressure by establishing subsets with higher survival potential.

Keywords. antimicrobial resistance; Iran; principle coordinates analysis; ST131; virotype; virulence genes; whole-genome sequencing.

Sequence type 131 (ST131), the currently emerged clone of *Escherichia coli* that is disseminated worldwide, causes severe hospital-acquired and community-onset infections [1, 2]. The pervasiveness of ST131 has been reported by many global surveys, and increasing prevalence of fluoroquinolone and cephalosporin resistance in patients with *E. coli* is attributed to this clone [3].

ST131 strains are closely related and appear to have had a common ancestor, so they are often referred to as a clone, or clonal, group [4]. However, whole-genome sequencing analysis of ST131 strains has revealed that this clone is not uniform, and 3 different clades, including clades A, B, and C, are characterized among this clone [5]. Generally, the clades A and B, which are minor parts of the ST131 population, are susceptible to fluoroquinolone and cephalosporin [6]. In contrast, clade C (also known as *H*30) represents the largest clade and comprises 2 subclades: C1 (or *H*30R) and C2 (*H*30Rx), both of which are resistant to fluoroquinolone [5]. Furthermore, phylogenetic tree analysis and carriage of a unique prophage-like region, have divided subclade C1 into 2 subsets, named C1-M27 and C1-non-M27 (C1-nM27) [7]. Apart from the extensively antibiotic-resistant phenotypes, which are identified among these strains, ST131 is also considered a highly virulent clone due to the higher capability of causing extraintestinal infections as compared with other clones [6]. This feature is attributed to the diverse putative virulence genes harbored by these strains [8].

Despite the highly conserved sequences that are identified in the core genome of ST131, the accessory genome of this clone is highly variable and results in differences in virulence

Received 14 May 2020; editorial decision 31 August 2020; accepted 2 September 2020.

Correspondence: O. Pajand, PhD, Semnan University of Medical Sciences, Semnan, Iran 3513138111 (om24pa@gmail.com).

Open Forum Infectious Diseases®

DOI: 10.1093/ofid/ofaa410

gene content and plasmid repertoire [9]. Considering virulence gene content, the ST131 clone can be categorized into 12 virotypes, which are named from A to F [6]. While virotype C has been reported to be the most common virotype among ST131 clones, the other virotypes do not have equal distribution among individuals with ST131 from different continents [10].

Detection of ST131 and its clades/subclades is important for epidemiological studies. While this has been recognized as a pandemic clonal group that threatens public health, ST131 has received less attention in Iran than other antimicrobial-resistant pathogens. Increasing rates of resistance against cephalosporins and fluoroquinolones among *E. coli* isolates from extaintestinal infections, mostly in a ST131 population, have been reported in recent years in Iran [11, 12]. In our current study, we aimed to identify the clades and virotypes of the ST131 population cultured from extraintestinal specimens during a 19-month surveillance study and determine the differences of antibiotic susceptibility patterns, virulence, and resistance markers between ST131 clades.

METHODS

Strains

In this 19-month cross-sectional study (March 2015 to September 2016), 338 *E. coli* isolates were cultured from patients with extraintestinal infections admitted to Kosar University Hospital in Semnan, Iran. Clinical samples were collected as part of standard care for admitted patients. Isolates were cultured from different specimens, including urine, blood, wound, and respiratory samples. The genomic DNA of isolates was extracted based on the Cetyl trimethylammonium bromide (CTAB) method [13]. Based on *gyrB*/*mdh* single nucleotide polymorphism (SNP) multiplex polymerase chain reaction (PCR) [14], 73 nonduplicate phylogroup B2 ST131 isolates were identified among this bacterial collection. So, the overall prevalence of detected ST131 strains was 21.5%. The O25b/O16 subgroups were determined as described earlier [15]. Allele-specific primers for allele 30 of *fimH* corresponding with the main fluoroquinolone resistance–associated subset within this clone were used to identify the *H*30 subclone [16].

Patient Consent Statement

Isolates were taken as part of routine hospital procedure; therefore, patient consent was not required. This study was approved by ethical committee of Semnan University of Medical Sciences with the ethics code IR.SEMUMS.REC.1398.219.

Clade Determination

For determination of ST131 clades, multiplex PCR using 7 pairs of primers was used as described by Matsumura et al. [17]. Clades and subclades were identified based on the expected amplicons. Amplification was performed using ready-to-use Master Mix (Tempase 2X Master Mix, Amplicon, Denmark) and recommended concentrations of primers [17].

Virulence Factors and Virotype Determination

The presence of 35 putative virulence markers was assessed by multiplex PCR [18–21]. Urinary pathogenic *E. coli* (UPEC) isolates were those strains that harbored ≥3 virulence genes, including *yfcV*, *fyuA*, *vat*, and *chuA*. The virulence factor (VF) score was the total number of virulence genes detected, adjusted for multiple detection of the *pap* operon [22]. The virotype of the ST131 isolates was established according to the scheme described by Dahbi et al. [14].

Antimicrobial Susceptibility Testing

The standard disk diffusion method on Mueller-Hinton agar was used to determine the antibiotic susceptibility patterns of 73 ST131 strains, and results were interpreted according to Clinical and Laboratory Standard Institute (CLSI) guidelines [23]. The number of antibiotics to which the strain was resistant was considered the resistance score. Isolates with resistance to at least 1 representative of 3 or more antimicrobial classes were defined as multidrug-resistant (MDR) [24]. Extended spectrum beta-lactamase (ESBL) production was assayed using phenotypic combined disk testing according to the recommendations of the CLSI [23].

Detection of Resistance Encoding Genes

The presence of resistance genes including ESBLs (bla_{TEM}, bla_{SHV}, bla_{OXA-1}, and $bla_{CTX-M-15}$) [25, 26] and plasmid-mediated quinolone resistance (PMQR; *qnrA*, *qnrB*, *qnrS*, and *aac(6')-Ib-cr*) was investigated by multiplex PCR according to previously published methods [27]. Furthermore, isolates harbouring the 16S *rRNA* methylase genes (*armA*, *rmtB*, *rmtC*) and aminoglycoside resistance determinants (*aac(3)-IIa*, *aac(6')-Ib*) were detected by single PCR [28, 29].

Statistical Analysis

To compare the proportions and scores, the Fisher exact test and Mann-Whitney *U* test were used, respectively. Principal coordinates analysis, a multidimensional scaling method analogous to principal component analysis, was used to collapse the molecular data set for simplified between-group comparisons [18]. Groups were compared on each of the first 3 coordinates, which captured most of the variance within the data set using a 2-tailed *t* test. *P* values <.05 were considered statistically significant.

RESULTS

Clade Determination, O25b/O16 Subgroups, and Virulence Gene Content

Multiplex PCR for clade determination revealed the C2 subclade as the dominant subset (33/73; 45.2%), followed by C1-M27 (18; 24.6%), C1-non-M27 (C1-nM27; 14; 19.1%) and

A (8; 10.9%). Strains of clade A were identified as belonging to the O16 subgroup and fim*H*30 negative, while the remaining 65 isolates belonged to the O25b subgroup and harboured the fim*H*30 allele. Seven virulence factors, including *sfa focDE*, *colV*, *ibeA*, *papGIII*, *cdtB*, and *neuCK*, were not detected among study isolates. The 29 virulence markers were detected at least once, with the lowest rate being 1.3% (*papA*, *papEF*, *hlyF*) and the highest rate being 100% (*usp, yfcV, fyuA, chuA*). Except for 2 C1 subsets (C1-M27, C1-nM27), the other 2 clades were considerably different in VF content, with the lowest VF score in clade A (median, 14) and the highest VF score in subclade C2 (median, 19).

Among the C1 strains, capsular type *kpsMT II* was significantly detected among C1-nM27, while *papC, papGII*, *hlyA*, *cnf1*, and *tia/hek* were negatively associated with both of the C1 subsets. Seven virulence markers (*iss*, *hra*, *cnf1*, *hlyA*, *iroN*, *papGII*, *papC*, and *tia/hek*) were significantly associated with subclade C2. Furthermore, a positive association was found between the carriage of *hlyA* and *papC*, *papGII*, *K5*, *hra*, *tia/hek*, and the *cnf1* gene within this population.

Principal coordinates analysis based on the 29 virulence determinants revealed that virulence profiles of subclade C2 were distinctive and differentiated from other strains. In Figure 1, a plotted coordinate 1–coordinate 2 plane, a variance of 58.77% was captured and subclade C2 isolates are clustered in the lower left quadrant, clearly separated from the other strains (Figure 1). The differences in aggregate virulence profiles were explored using univariate analysis. Subclade C2 strains showed higher aggregate virulence scores (median) than other clades. All isolates fulfilled the molecular criteria for UPEC. Table 1 shows the prevalence of virulence genes among different clades. As part of the other research project, studied isolates were subjected to whole-genome sequencing (WGS). The PCR results of virulence factors were confirmed based on WGS results obtained from analyzing the assembled draft genomes using the VFDB

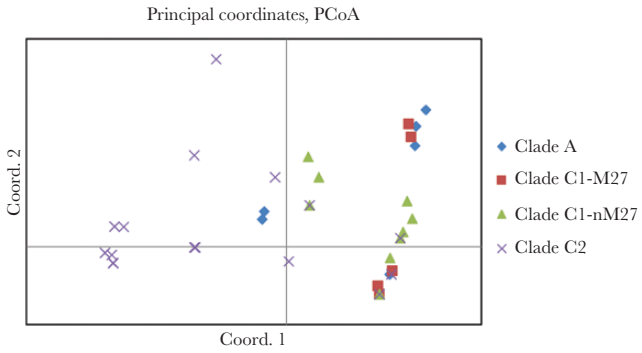

Figure 1. Principal coordinate analysis (PCoA) of virulence gene profiles among 73 ST131 isolates. The PCoA was based on results for all 29 virulence genes studied. Each isolate is plotted based on its values for PCoA coordinates 1 (x-axis) and 2 (y-axis), which collectively capture 58.77% of the total variance in the data set.

Table 1. Prevalence of Virulence Markers Among ST131 Clades

	A	C1-M27	C1-nM27	C2	
	n = 8 (10.9%)	n = 18 (24.6%)	n = 14 (19.1%)	n = 33 (45.2%)	Total n = 73
Virulence Factors			No. (%)		
Source of isolates					
	Sputum: 1 UC: 7	Sputum: 2 UC: 16	Sputum: 4 UC: 8 Wound: 2	UC: 31 Wound: 2	
Adherence					
papA	1 (12.5)	0	0	0	1 (1.3)
P value					
papC	2 (25)	0	0	27 (81.8)	29 (39.7)
P value		*<.001*	*<.001*	**<.001**	
papGII	2 (25)	0	0	28 (78.8)	30 (41.09)
P value		*<.001*	*.001*	**<.001**	
papEF	0	0	*0*	*1 (3)*	1 (1.3)
F10papA	8 (100)	16 (88.9)	14 (100)	33 (100)	73 (100)
afaDrBC	4 (50)	0	2 (14.2)	3 (9.1)	9 (12.3)
P value	**.006**	*.05*			
afaFM955459	2 (25)	0	2 (14.2)	3 (9.1)	7 (9.5)
iha	7 (87.5)	16 (88.9)	13 (92.9)	32 (97)	68 (93.1)
hra	0	0	3 (21.4)	29 (87.9)	32 (43.8)
P value[a]	*.008*	*<.001*		**<.001**	
Iron uptake					
iucdD	5 (62.5)	16 (88.9)	14 (100)	29 (87.9)	63 (87.6)
P value					
iutA	5 (62.5)	16 (88.9)	13 (92.9)	31 (93.9)	65 (89)
P value	*.03*				
yfcV	8 (100)	18 (100)	9 (100)	33 (100)	73 (100)
fyuA	8 (100)	18 (100)	9 (100)	33 (100)	73 (100)
chuA	8 (100)	18 (100)	9 (100)	33 (100)	73 (100)
iroN	0	0	0	5 (15.2)	5 (6.8)
				0.01	
Invasion					
Tia/hek	2 (25)	0	3 (21.4)	29 (87.9)	34 (46.5)
P value		*<.001*	*.04*	**<.001**	
Autotransporter					
sat	5 (62.5)	16 (88.9)	13 (92.9)	32 (97)	66 (90.4)
P value	.02				
vat	2 (25)	0	0	0	2 (2.7)
P value	.01				
tsh	7 (87.5)	18 (100)	14 (100)	33 (100)	72 (98.6)
Toxins					
cnF1	0	0	0	21 (61.6)	21 (28.7)
P value		*.001*	*.007*	**<.001**	
hlyA	0	0	0	21 (61.6)	21 (28.7)
P value		*.001*	*.007*	**<.001**	
usp	8 (100)	18 (100)	9 (100)	33 (100)	73 (100)
hlyF	0	0	0	1 (3)	1 (1.3)
Protection					
kpsMTII	1 (12.5)	0	7 (50)	3 (9.1)	11 (15)
P value		*.05*	**<.001**		
K5	7 (87.5)	17 (94.4)	5 (35.7	28 (84.8)	57 (78)
P value			*<.001*		
Iss	0	16 (88.9)	13 (92.9)	33 (100)	62 (84.9)
	<0.001			**0.001**	
traT	8 (100)	17 (94.4)	12 (85.7)	31 (93.9)	68 (93.1)

Table 1. Continued

	A	C1-M27	C1-nM27	C2	
	n = 8 (10.9%)	n = 18 (24.6%)	n = 14 (19.1%)	n = 33 (45.2%)	Total n = 73
Virulence Factors			No. (%)		
Miscellaneous					
PAI	6 (75)	18 (100)	13 (92.9)	33 (100)	70 (95.8)
P value	*.03*				
ompT	7 (87.5)	18 (100)	14 (100)	33 (100)	92 (98.6)
VF score (Mean, Median)	13.75, 14	14.28, 15	15.14, 15	18.6, 19	

Abbreviation: UC, urine culture.

[a]Comparison between each clade and all other clades combined. Values in boldface indicate significant associations. *P* values are shown for differences that were statistically significant ($P < .05$). Italic formatting indicates a negative association.

[30] and VirulenceFinder 2.0 virulence gene databases (data not shown) [31].

Clades and Resistance Profiles/Resistance Determinants

Clade A was significantly associated with susceptibility to fluoroquinolone. The 2 C1 subclades showed different patterns of antibiotic susceptibility profiles. C1-M27 was significantly associated with susceptibility to ampicillin/sulbactam, amoxicillin/clavulanate, gentamicin, and tobramycin, and all were phenotypically ESBL/MDR; however, no such association was found for the C1-nM27 subclade. The resistance rates were increased from clade A with median of 3 to subclade C2 (median, 5). Subclade C2 showed significantly higher rates of resistance against gentamicin, amikacin, tobramycin, amoxicillin-clavulanate, ampicillin-sulbactam, nitrofurantoin, and fluoroquinolone. The most frequently detected resistance marker was $bla_{\text{CTX-M-15}}$, which was identified among all strains (45; 61.6%) except subclade C1-M27. Of the resistance markers studied, $bla_{\text{CTX-M-15}}$, *aac(3)-IIa*, *aac(6')-Ib*, *aac(6')-Ib-cr*, and $bla_{\text{OXA-1}}$ were associated with subclade C2, while this clade was conspicuous for a low prevalence of bla_{TEM} ($P = .01$) (Table 2). The PCR results of resistance genes were confirmed by analyzing assembled draft genomes obtained from WGS of isolates

Table 2. Antibiotic Resistance Rates and Prevalence of Resistance Markers Among Different Clades of ST131

	A	C1-M27	C1-nM27	C2	Total
	n = 8	n = 18	n = 14	n = 33	(%)
Antibiotics			No. (%)		
Imipenem	1 (12.5)	(0)	1 (7.1)	0	2 (2.7)
Meropenem	0	1 (5.6)	0	0	1 (1.4)
Ertapenem	1 (12.5)	2 (11.1)	2 (14.3)	4 (12.1)	9 (12.3)
Piperacillin-tazobactam	2 (25)	3 (16.7)	2 (14.3)	7 (21.2)	14 (19.2)

Table 2. Continued

	A	C1-M27	C1-nM27	C2	Total
	n = 8	n = 18	n = 14	n = 33	(%)
Antibiotics			No. (%)		
Ampicillin-sulbactam	5 (62.5)	2 (11.1)	5 (35.7)	20 (60.6)	32 (43.8)
P value[a]		*.002*		**.01**	
Amoxicillin-clavulanate	6 (75)	5 (27.8)	7 (50)	26 (78.8)	44 (60.3)
P value		*.002*		**.004**	
Trimethoprim/sulfamethoxazole	6 (75)	14 (77.8)	8 (57.1)	23 (69.7)	51 (69.9)
Aztreonam	6 (75)	16 (88.9)	12 (85.7)	31 (93.9)	65 (89)
Cefepime	5 (62.5)	8 (44.4)	7 (50)	22 (66.7)	42 (57.5)
Ceftazidime	6 (75)	12 (66.7)	11 (78.6)	27 (81.8)	56 (76.5)
Cefotaxime	6 (75)	18 (100)	12 (85.7)	32 (97)	68 (93.2)
Amikacin	0	0	0	8 (24.2)	8 (11)
P value				**.001**	
Gentamicin	2 (25)	1 (5.6)	3 (21.4)	25 (75.8)	31 (42.5)
P value		*<.001*		**<.001**	
Tobramycin	1 (12.5)	0	2 (14.3)	30 (90.2)	33 (45.2)
P value		*<.001*	*.01*	**<.001**	
Ciprofloxacin	0	18 (100)	13 (92.9)	33 (100)	64 (87.7)
P value	*<.001*			**.003**	
Levofloxacin	0	18 (100)	14 (100)	33 (100)	65 (89)
P value	*<.001*			**.007**	
Nitrofurantoin	0	1 (5.6)	1 (7.1)	8 (24.2)	10 (13.7)
P value				**.03**	
MDR	6 (75)	18 (100)	10 (71.4)	33 (97.1)	69 (90.8)
			0.02		
Resistance rates (Mean, Median)	2.63, 3	3.39, 3	3.36, 3.50	4.74, 5	
ESBL	6 (75)	18 (100)	12 (85.7)	33 (97.1)	69 (94.5)
Resistance markers					
$bla_{\text{CTX-M-15}}$	5 (62.5)	0	9 (64.3)	31 (93.9)	45 (61.6)
P value		*<.001*		**<.001**	
$bla_{\text{TEM-}}$	8 (100)	0	9 (64.3)	7 (21.2)	24 (32.8)
P value	**<.001**	*<.001*	**.01**	*.01*	
$bla_{\text{OXA-1}}$	0	0	0	29 (87.9)	29 (39.7)
P value	*.01*	*<.001*	*<.001*	**<.001**	
aac6Ib-cr	0	0	1 (7.1)	27 (81.8)	28 (38.3)
P value	*.02*	*<.001*	*<.01*	**.001**	
qnrS	0	0	4 (28.6)	0	4 (5.4)
P value			**.001**		
aac3IIa	6 (75)	0	4 (28.6)	23 (69.7)	33 (45.2)
P value		*<.001*		**<.001**	
aac6Ib	1 (12.5)	0	1 (7.1)	28 (84.8)	30 (41)
P value		*<.001*	*.005*	**<.001**	

Abbreviation: ESBL, extended spectrum beta-lactamase.

[a]Comparison between each clade and all other clades combined. Bold values indicate significant associations. *P* values are shown for differences that were statistically significant ($P < .05$). Italic formatting indicates a negative association.

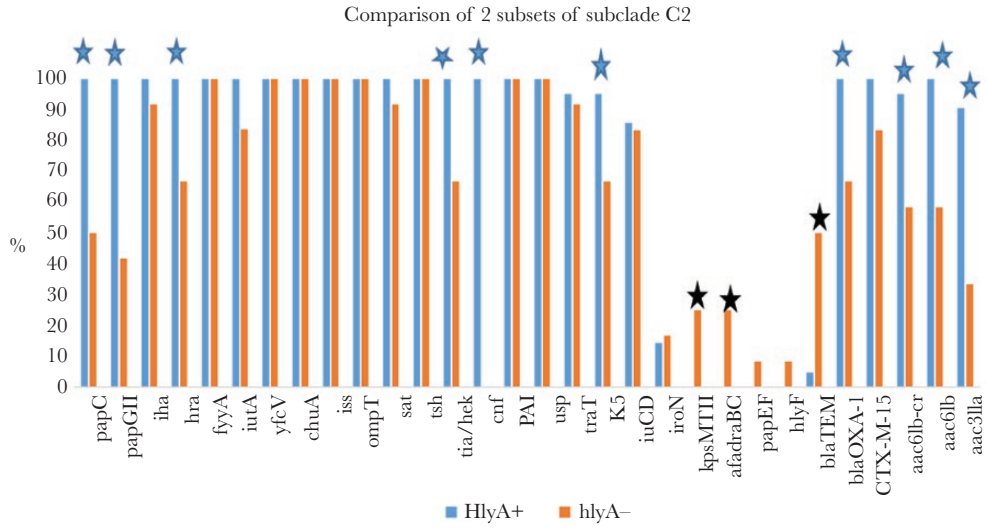

Figure 2. Comparison of virulence/resistance genes content between 2 subsets (*hly A*$^+$ and *hlyA*$^-$) of subclade C2 strains. Black and blue stars indicate the association of that gene with *hlyA*$^-$ and *hlyA*$^+$ strains, respectively.

using the ResFinder antimicrobial resistance gene database (data not shown) [32].

Closer examination of the C2 subclade showed that the *hlyA* gene was not uniformly distributed among this clade and strains carrying this virulence marker were significantly positive for *aac(6')-Ib* (21; 100%; P = .003), *aac(6')-Ib-cr* (20; 95.2%; P = .01), $bla_{\text{OXA-1}}$ (21; 100%; P = .01), and *aac(3)-IIa* (19; 90.5%; P = .001) (Figure 2). The carriage of *hlyA* among the C2 subclade coincided with the resistance phenotype to tobramycin (21; 100%; P = .04) and gentamicin (21; 100%; P < .001). In total, the $bla_{\text{CTX-M-15}}$-positive isolates exhibited a significantly higher prevalence of resistance to aztreonam (45; 100%; P < .001), ceftazidime (39; 86.7%; P = .02), cefepime (34; 75.6%; P < .001), ampicillin/sulbactam (27; 60%; P = .001), amoxicillin/clavulanate (33; 73.3%; P = .006), amikacin (8; 17.8%; P = .02), tobramycin (32; 71.1%; P < .001), and gentamicin (29; 64.4%; P < .001).

Considering $bla_{\text{OXA-1}}$, there was a strong association between the carriage of this element and resistance to amikacin, gentamicin, tobramycin, aztreonam, amoxicillin/clavulanate, and ampicillin/sulbactam and fluoroquinolone.

Virotyping of Clades and Resistance Markers

All except 13 strains were divided into virotypes A to F. These 13 strains showed unknown arrangements of virulence genes, and their virulence patterns are shown in Table 3.

Table 3. Unknown Virotypes Detected among 13 Strains. Clades and Serotypes for Each Pattern are Shown.

	Virulence Genes[a]													
Unknown Virotypes n = 13	*afaDraBC*	*afaFM955459*	*iroN*	*sat*	*papGII*	*cnf*	*hlyA*	*kpsMTII*	*k5*	*papC*[b]	*vat*[b]	*iss*[b]	Clades	O25/O16
Pattern 1 n = 2	+	+	-	+	+	-	-	-	+	+	+	-	A	O16
Pattern 2 n = 2	+	+	-	+	-	-	-	-	+	-	-	+	C2	O25
Pattern 3 n = 2	+	-	-	-	-	-	-	-	+	-	-	-	A	O16
Pattern 4 n = 3	-	-	+	+	+	+	+	-	+	+	-	+	C2	O25
Pattern 5 n = 1	-	-	-	-	-	-	-	+	-	-	-	-	A	O16
Pattern 6 n = 1	-	-	-	-	-	-	-	-	+	-	-	-	C1-M27	O25
Pattern 7 n = 2	-	-	-	-	-	-	-	-	-	-	-	+	C1	O25

[a]The virulence genes *cdtB, neuC k, papGIII,* and *ibeA,* which are used for virotyping, were not detected.

[b]These 3 genes are not considered for virotyping and are shown because of the same arrangement among isolates of each pattern.

Table 4. Virotyping of Clades and Prevalence of Resistance Genes Among Virotypes

	Virotype A	Virotype B	Virotype C	Virotype E	Virotype F	Unknown
	(n = 3)	(n = 2)	(n = 34)	(n = 18)	(n = 3)	(n = 13)
Clades						
Resistance markers	C1-nM27: 2 C2: 1	C2: 2	A: 3 C1-M27: 16 C1-nM27: 11 C2: 4	C2: 18	C2: 3	A: 5 C1-M27: 2 C1-nM27: 1 C2: 5
bla_{TEM}	3 (100)	1 (50)	11 (32.4)	1 (5.6)	2 (66.7)	6 (46.2)
P value				*.004*		
bla_{OXA-1}	1 (33.1)	1 (50)	2 (5.9)	18 (100)	3 (100)	4 (30.8)
P value				**<.001**		
$bla_{CTX-M-15}$	1 (33.3)	1 (50)	13 (38.2)	18 (100)	3 (100)	9 (69.2)
P value				**<.001**		
aac6Ib	1 (33.3)	1 (50)	4 (11.8)	18 (100)	2 (66.7)	4 (30.8)
P value				**<.001**		
aac3IIa	0	0	7 (20.6)	17 (94.4)	2 (66.7)	3 (23.1)
P value				**<.001**		
aac6Ib-cr	1 (33.3)	1 (50)	3 (8.8)	17 (94.4)	2 (66.7)	4 (30.8)
P value				**<.001**		
qnrS			4 (11.8)			
Virulence score (Mean, Median)	17.67, 17	14.50, 14.50	14.94, 15	18.83, 19	17	15.17, 16
Resistance score (Mean, Median)		3.50, 3.50	3.62, 4	4.94, 5		3.17, 3

[a]Thirteen strains were not categorized as defined virotypes. Bold values indicate a significant association. Italic formatting indicates a negative association.

Virotype C was the most common virotype, represented by 34 (46.5%) isolates and predominantly associated with subclade C1 (27/32), including the C1-M27 (16 strains within virotype C; 47%) and C1-nM27 (11 strains within virotype C; 32.3%) subsets. Most of the virotype C strains were identified as virotype C2 (25; 73.5%), followed by virotype C1 (7; 20.5%) and virotype C3 (2; 6%). In contrast, virotypes E (18 strains) and F (3 strains) belonged to subclade C2 (Table 4).

The associations among all 29 detected virulence genes were identified using cluster analysis. The isolates were divided according to 100% similarity of virulence gene content. The largest cluster included 20 strains with a unique set of 15 virulence genes (*sat*, *chuA*, *fyuA*, *yfcV*, *iutA, k5, iucD*, *F10papA*, *iha*, *tsh*, *ompT*, *usp*, *PAI*, *traT*, *iss*) that corresponded to virotype C. These isolates harbored mostly $bla_{CTX-M-15}$ (6/20; 30.%), or were bla_{OXA-1} negative, and mainly represented the C1 subclade (19/20; 95%; including 14 C1-M27 and 5 C1-nM27 strains).

The second largest cluster, which included 14 subclade C2 strains, contained bla_{OXA-1} (100%), $bla_{CTX-M-15}$ (100%), *aac(3)-IIa* (92.8%), *aac(6')-Ib* (100%),and *aac(6')-Ib-cr* (92.8%) positive strains with a set of 21 virulence markers (*papGII*, *papC*, *F10papA*, *iucD*, *sat*, *cnf1*, *hlyA*, *chuA*, *fyuA*, *yfcV*, *iutA*, *iha*, *hra*, *tsh*, *tia/hek*, *usp, K5, ompT, iss*, *traT*, and *PAI*) that corresponded to virotype E. As expected, these 14 virotype E strains were significantly associated with resistance phenotypes to gentamicin ($P < .001$) and tobramycin ($P < .001$).

The highest VF score was detected in virotype E (median, 19), followed by virotype F (median, 17), virotype A (median, 17), and unknown virotype (median, 16).

DISCUSSION

As far as we know, this is the first study in Iran to investigate and compare the prevalence and genotypes of ST131 subclades. Here, we found that the C2 subclade of ST131 was responsible for most of the ST131 infections. All except clade B were detected in the study population. Interestingly, virotype and clade patterns had consistency, in which the most common virotypes including virotypes C and E comprised the subclade C1 and C2 strains, respectively. A subpopulation among the subclade C2 lineage was detected that showed higher carriage rates of resistance/virulence markers and was particularly resistant to aminoglycosides, confirming the importance of emerged subsets within this clone.

In our survey, an increasing trend in the resistance rate was detected from C1-M27 to C1-nM27 and the C2 strains. Considering the susceptibility patterns, significant differences were observed between subclades, particularly between C1-M27 and C2, despite their phylogenetic relatedness. These differences were remarkable in the proportion of resistance to ampicillin/sulbactam, amoxicillin/clavulanate, gentamicin, and tobramycin, which was very low among C1-M27 strains compared with the other isolates. Furthermore, C1-M27 was conspicuously negative for the studied resistance markers. This finding indicates that we need to investigate factors other than antibiotic selection pressure to explain the emergence of C1-M27 as the second most prevalent subclade. In a 4-year study on isolated *E. coli* strains from blood cultures in Norway, the resistance rate against gentamicin was found to be much higher than that of cephalosporin among O25b-ST131 isolates

[33]. In fact, the appearance of a subpopulation with remarkable resistance to aminoglycosides among ST131, which is notorious for fluoroquinolone resistance and producing ESBL, represents an evolving epidemiology of this clone and its subclades to a pandrug-resistant population [34].

Here, we identified 7 new virotypes including 13 strains with different genetic arrangements, suggesting an endemic distribution of virulence markers probably acquired by mobile genetic elements. Among the known virotypes, virotype C is considered to be the most widely distributed virotype in ST131, occurring in all ST131 clades [35]. Here, virotype C was predominant in subclade C1, specifically the C1-M27 strains, as also reported recently from Southwest England and Europe [35, 36]. In contrast to several European studies that reported virotype A as a dominant virotype among subclade C2 [10], virotypes E and F were found to be the most common virotypes, representative of >60% of the C2 strains (21 out of 33 strains). Further analysis of subclade C2 revealed a heterogeneous population based on the carriage of the *hlyA* virulence marker, as positive strains had remarkably higher rates of some virulence (*hra*, *tia/hek*, *cnf*, *papC*, *papGII*, *K5*) and all except bla_{TEM} resistance genes, and consequently a higher resistance phenotype to tobramycin and gentamicin. Interestingly, almost all of this C2 subclade (18 *hlyA*+/21 *hlyA*+) was identified as virotype E. In recently published data from Southeast Asia, a subpopulation among the C2 subclade was reported that has been named Southeast Asia-C2 (SEA-C2) lineage [37]. While virotypes of studied strains were not determined in the aforementioned study, the main features attributed to this subset were the higher carriage rates of some virulence genes, mainly *hly*ABCD operon, *cnf1*, and *tia/hek*, and harboring a conserved plasmid that carried *aac(3)-IIa*, *aac(6')-Ib, aac(6')-Ib-cr*, $bla_{CTX\text{-}M\text{-}15}$, $bla_{OXA\text{-}1}$, and *tetA*. The strains were very closely related by genome sequence analysis. These findings in conjunction with our data suggest that virotypes E and F are the most prevalent virotypes among the C2 subclade originating from Asia, constituting a distinct subset among the C2 population.

Our study has some limitations, including a small sample size that was collected from a single center, lack of knowledge regarding whether the study isolates were part of a nosocomial outbreak, and, most importantly, a lack of data regarding phylogenic clustering of strains based on a roboust technique such as whole-genome sequencing.

In conclusion, our study is notable for examining 19 months' worth of collected ST131 strains in a geographical region from which no data have been previously published. We found that ST131 strains are not a uniform population and that subclade C2, like in other regions, drives both the higher virulence and resistance among this high-risk clone. Greater focus on this clade identified a subset of strains that showed virotype E and F patterns, and their resistance phenotypes and resistance/virulence gene repertoires were different from the other C2 subclade strains. So, our data show a local trend within the C2 subclade in which the generated subset has a major advantage over other individuals with ST131 in the context of resistance/virulence gene content and clonality. So, ongoing monitoring of the dynamics of ST131 local transmission is required to understand the reasons for new virotypes' emergence within this region.

Acknowledgments

With a broken heart, I write this tribute in loving memory of Dr Zoya Hojabri, my wife and the first author of this article. On September 5th, 2020 both Zoya and our two-year-old son, Bardiya, died in a tragic car accident.

Zoya earned her PhD in the field of Medical Microbiology and Epidemiology at Tabriz University in 2013; and in 2014 when she was just 30 years old, she began her professorship at the Semnan University of Medical Sciences. Committed to improving the fundamentals of the teaching system, Zoya was a passionate teacher who cared deeply for her students, focused on cultivating critical thinking skills in budding researchers, especially for young women in science. She devoted herself to the development of microbiology research laboratory, designed specifically for the study of molecular epidemiology of Carbapenem-resistant Enterobacteriaceae (CRE), and whole-genome sequencing for food microbiology – the first of its kind, breaking new ground in Iran. Though her scientific career spanned less than fifteen years, Zoya published over 25 peer-reviewed articles in prestigious scientific journals and was well on her way to a distinguished academic career.

Outside the laboratory, Zoya loved and embraced life to the fullest; she cherished nature and her greatest joy was hiking in the mountains. When Bardiya was born, she became a new woman, and within herself found a strength I will forever admire. They were attached at the hip, yet she managed her duties of the professorship, of setting up and running an advanced research laboratory, and mentored countless medical students along the way. Even when I was away for a brief research sabbatical in Denmark, she single-handedly worked with the vigor of multiple men, and did it all with a beaming smile on her face; she had found her calling as a mother and a scientist. I owe a tremendous gratitude to this woman, my wife, my rock, my colleague, and my only love.

The loss of both Zoya and Bardiya has left me with a tremendous void – but I will continue this work in her honour and will work to the last of my days to honour her legacy and devotion to science. Zoya, Bardiya, dar ghalbe manid – I will love you forever.

Dr. Omid Pajand

October 2020

Financial support. This work was supported by Semnan University of Medical Sciences (grant Nos. 730 and 1636).

Potential conflicts of interest. The authors declare that they have no competing interests. All authors: no reported conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Nicolas-Chanoine MH, Blanco J, Leflon-Guibout V, et al. Intercontinental emergence of *Escherichia coli* clone O25:H4-ST131 producing CTX-M-15. J Antimicrob Chemother **2008**; 61:273–81.
2. Mendis SM, Vasoo S, Johnston BD, et al. Clinical and molecular correlates of *Escherichia coli* bloodstream infection from two geographically diverse centers in Rochester, Minnesota, and Singapore. Antimicrob Agents Chemother **2018**; 62:e00937–18.
3. Nicolas-Chanoine MH, Bertrand X, Madec JY. *Escherichia coli* ST131, an intriguing clonal group. Clin Microbiol Rev **2014**; 27:543–74.
4. Shaik S, Ranjan A, Tiwari SK, et al. Comparative genomic analysis of globally dominant ST131 clone with other epidemiologically successful extraintestinal pathogenic *Escherichia coli* (ExPEC) lineages. MBio **2017**; 8:5.
5. Forde BM, Roberts LW, Phan MD, et al. Population dynamics of an *Escherichia coli* ST131 lineage during recurrent urinary tract infection. Nat Commun **2019**; 10:3643.
6. Mamani R, Flament-Simon SC, García V, et al. Sequence types, clonotypes, serotypes, and virotypes of extended-spectrum β-lactamase-producing *Escherichia coli* causing bacteraemia in a Spanish hospital over a 12-year period (2000 to 2011). Front Microbiol **2019**; 10:1530.
7. Matsumura Y, Pitout JD, Gomi R, et al. Global *Escherichia coli* sequence type 131 clade with $bla_{\text{CTX-M-27}}$ gene. Emerg Infect Dis **2016**; 22:1900–07.
8. Merino I, Porter SB, Johnston B, et al. Molecularly defined extraintestinal pathogenic *Escherichia coli* status predicts virulence in a murine sepsis model better than does virotype, individual virulence genes, or clonal subset among *E. coli* ST131 isolates. Virulence **2020**; 11:327–36.
9. Petty NK, Ben Zakour NL, Stanton-Cook M, et al. Global dissemination of a multidrug resistant *Escherichia coli* clone. Proc Natl Acad Sci U S A **2014**; 111:5694–9.
10. Jamborova I, Johnston BD, Papousek I, et al. Extensive genetic commonality among wildlife, wastewater, community, and nosocomial isolates of *Escherichia coli* sequence type 131 (*H*30R1 and *H*30Rx subclones) that carry $bla_{\text{CTX-M-27}}$ or $bla_{\text{CTX-M-15}}$. Antimicrob Agents Chemother **2018**; 62:e00519–18.
11. Neamati F, Khorshidi A, Moniri R, Hosseini Tafreshi SA. Molecular epidemiology of antimicrobial resistance of uropathogenic *Escherichia coli* isolates from patients with urinary tract infections in a tertiary teaching hospital in Iran. Microb Drug Resist **2020**; 26:60–70.
12. Jafari A, Falahatkar S, Delpasand K, et al. Emergence of *Escherichia coli* ST131 causing urinary tract infection in Western Asia: a systematic review and meta-analysis. Microb Drug Resist. **In press**.
13. Hojabri Z, Arab M, Darabi N, et al. Evaluation of the commercial combined disk test and minimum inhibitory concentration (MIC) determination for detection of carbapenemase producers among gram-negative bacilli isolated in a region with high prevalence of $bla_{\text{OXA-48}}$ and bla_{NDM}. Int Microbiol **2019**; 22:81–9.
14. Dahbi G, Mora A, Mamani R, et al. Molecular epidemiology and virulence of *Escherichia coli* O16:H5-ST131: comparison with H30 and H30-Rx subclones of O25b:H4-ST131. Int J Med Microbiol **2014**; 304:1247–57.
15. Johnson JR, Clermont O, Johnston B, et al. Rapid and specific detection, molecular epidemiology, and experimental virulence of the O16 subgroup within *Escherichia coli* sequence type 131. J Clin Microbiol **2014**; 52:1358–65.
16. Colpan A, Johnston B, Porter S, et al; VICTORY (Veterans Influence of Clonal Types on Resistance: Year 2011) Investigators. *Escherichia coli* sequence type 131 (ST131) subclone H30 as an emergent multidrug-resistant pathogen among US veterans. Clin Infect Dis **2013**; 57:1256–65.
17. Matsumura Y, Pitout JDD, Peirano G, et al. Rapid identification of different *Escherichia coli* sequence type 131 clades. Antimicrob Agents Chemother **2017**; 61.
18. Hojabri Z, Mirmohammadkhani M, Darabi N, et al. Characterization of antibiotic-susceptibility patterns and virulence genes of five major sequence types of *Escherichia coli* isolates cultured from extraintestinal specimens: a 1-year surveillance study from Iran. Infect Drug Resist **2019**; 12:893–903.
19. Johnson JR, Porter S, Thuras P, et al. The pandemic *H*30 subclone of sequence type 131 (ST131) as the leading cause of multidrug-resistant *Escherichia coli* infections in the United States (2011–2012). Open Forum Infect Dis **2017**; 4:ofx089.
20. Ewers C, Janssen T, Kiessling S, et al. Molecular epidemiology of avian pathogenic *Escherichia coli* (APEC) isolated from colisepticemia in poultry. Vet Microbiol **2004**; 104:91–101.
21. Ewers C, Li G, Wilking H, et al. Avian pathogenic, uropathogenic, and newborn meningitis-causing *Escherichia coli*: how closely related are they? Int J Med Microbiol **2007**; 297:163–76.
22. Johnson TJ, Wannemuehler Y, Doetkott C, et al. Identification of minimal predictors of avian pathogenic *Escherichia coli* virulence for use as a rapid diagnostic tool. J Clin Microbiol **2008**; 46:3987–96.
23. Clinical and Laboratory Standards Institute. Performance Standards for Antimicrobial Susceptibility Testing; Twentieth Informational Supplement M100S-S26. Wayne, PA: CLSI; **2016**.
24. Johnson JR, Johnston B, Clabots C, et al. *Escherichia coli* sequence type ST131 as an emerging fluoroquinolone-resistant uropathogen among renal transplant recipients. Antimicrob Agents Chemother **2010**; 54:546–50.
25. Hojabri Z, Darabi N, Arab M, et al. Clonal diversity, virulence genes content and subclone status of *Escherichia coli* sequence type 131: comparative analysis of *E. coli* ST131 and non-ST131 isolates from Iran. BMC Microbiol **2019**; 19:117.
26. Blanco J, Mora A, Mamani R, et al; Spanish Group for Nosocomial Infections (GEIH). Four main virotypes among extended-spectrum-β-lactamase-producing isolates of *Escherichia coli* O25b:H4-B2-ST131: bacterial, epidemiological, and clinical characteristics. J Clin Microbiol **2013**; 51:3358–67.
27. Pajand O, Darabi N, Arab M, et al. The emergence of the hypervirulent *Klebsiella pneumoniae* (hvKp) strains among circulating clonal complex 147 (CC147) harbouring $bla_{\text{NDM/OXA-48}}$ carbapenemases in a tertiary care center of Iran. Ann Clin Microbiol Antimicrob **2020**; 19:12.
28. Hojabri Z, Mirmohammadkhani M, Kamali F, et al. Molecular epidemiology of *Escherichia coli* sequence type 131 and its H30/H30-Rx subclones recovered from extra-intestinal infections: first report of OXA-48 producing ST131 clone from Iran. Eur J Clin Microbiol Infect Dis **2017**; 36:1859–66.
29. Pajand O, Hojabri Z, Nahaei MR, et al. In vitro activities of tetracyclines against different clones of multidrug-resistant *Acinetobacter baumannii* isolates from two Iranian hospitals. Int J Antimicrob Agents **2014**; 43:476–8.
30. Chen L, Zheng D, Liu B, et al. VFDB 2016: hierarchical and refined dataset for big data analysis—10 years on. Nucleic Acids Res **2016**; 44:D694–7.
31. Joensen KG, Scheutz F, Lund O, et al. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. J Clin Microbiol **2014**; 52:1501–10.
32. Zankari E, Hasman H, Cosentino S, et al. Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother **2012**; 67:2640–4.
33. Fladberg ØA, Jørgensen SB, Aamot HV. Genotypic characterization of gentamicin and cephalosporin resistant *Escherichia coli* isolates from blood cultures in a Norwegian university hospital 2011–2015. Antimicrob Resist Infect Control **2017**; 6:121.
34. Merino I, Hernández-García M, Turrientes MC, et al; R-GNOSIS Study Group. Emergence of ESBL-producing *Escherichia coli* ST131-C1-M27 clade colonizing patients in Europe. J Antimicrob Chemother **2018**; 73:2973–80.
35. Findlay J, Gould VC, North P, et al. Characterization of cefotaxime-resistant urinary *Escherichia coli* from primary care in South-West England 2017–18. J Antimicrob Chemother **2019**; 75:65–71.
36. Olesen B, Frimodt-Møller J, Leihof RF, et al. Temporal trends in antimicrobial resistance and virulence-associated traits within the *Escherichia coli* sequence type 131 clonal group and its H30 and H30-Rx subclones, 1968 to 2012. Antimicrob Agents Chemother **2014**; 58:6886–95.
37. Chen SL, Ding Y, Apisarnthanarak A, et al. The higher prevalence of extended spectrum beta-lactamases among *Escherichia coli* ST131 in Southeast Asia is driven by expansion of a single, locally prevalent subclone. Sci Rep **2019**; 9:13245.