UWL REPOSITORY

repository.uwl.ac.uk

Short- and Long-Read Sequencing Reveals the Presence and Evolution of an IncF Plasmid Harboring blaCTX-M-15 and blaCTX-M-27 Genes in Escherichia coli ST131

This is a University of West London scholarly output.

Contact **open.research@uwl.ac.uk** if you have any queries.

Alternative formats: If you require this document in an alternative format, please contact: open.access@uwl.ac.uk

Short- and Long-Read Sequencing Reveals the Presence and Evolution of an IncF Plasmid Harboring $bla_{CTX-M-15}$ and $bla_{CTX-M-27}$ Genes in *Escherichia coli* ST131

Rory Cave,[a] Mary M. Ter-Stepanyan,[b,c] ⓘ Hermine V. Mkrtchyan[a]

[a]School of Biomedical Sciences, University of West London, London, United Kingdom
[b]Yerevan State Medical University after M. Heratsi, Faculty of Public Health, Department of Epidemiology, Yerevan, Republic of Armenia
[c]Research Center of Maternal and Child Health Protection, Yerevan, Armenia

ABSTRACT *Escherichia coli* sequence type 131 (ST131) has contributed to the spread of extended-spectrum beta-lactamase (ESBL) and has emerged as the dominant cause of hospital- and community-acquired urinary tract infections. Here, we report for the first time an in-depth analysis of whole-genome sequencing (WGS) of 4 ESBL-producing *E. coli* ST131 isolates recovered from patients in two hospitals in Armenia using Illumina short-read sequencing for accurate base calling to determine their genotype and to infer their phylogeny and using Oxford Nanopore Technologies long-read sequencing to resolve plasmid and chromosomal genetic elements. Genotypically, the four Armenian isolates were identified as part of the *H*30Rx/clade C2 ($n = 2$) and *H*41/clade A ($n = 2$) lineages and were phylogenetically closely related to isolates from the European Nucleotide Archive (ENA) database previously recovered from patients in the United States, Australia, and New Zealand. The Armenian isolates recovered in this study had chromosomal integration of the $bla_{CTX-M-15}$ gene in the *H*30Rx isolates and a high number of virulence genes found in the *H*41 isolates associated with the carriage of a rare genomic island (in the context of *E. coli* ST131) containing the S fimbrial, salmochelin siderophore, and microcin H47 virulence genes. Furthermore, our data show the evolution of the IncF[2:A2:B20] plasmid harboring both $bla_{CTX-M-15}$ and $bla_{CTX-M-27}$ genes, derived from the recombination of genes from an IncF [F2:A−:B−] $bla_{CTX-M-15}$-associated plasmid into the IncF[F1:A2:B20] $bla_{CTX-M-27}$-associated plasmid backbone seen in two genetically closely related *H*41 Armenian isolates.

IMPORTANCE Combining short and long reads from whole-genome sequencing analysis provided a genetic context for uncommon genes of clinical importance to better understand transmission and evolutionary features of ESBL-producing uropathogenic *E. coli* (UPEC) ST131 isolates recovered in Armenia. Using hybrid genome assembly in countries lacking genomic surveillance studies can inform us about new lineages not seen in other countries with genes encoding high virulence and antibiotic resistance harbored on mobile genetic elements.

KEYWORDS *E. coli* ST131, ESBL UPEC, hybrid assembly, long-read sequencing, short-read sequencing, whole-genome sequencing

*E*scherichia coli sequence type 131 (ST131) is the most dominant global lineage of *E. coli* to cause community and hospital-acquired urinary tract infections (UTI) (1, 2). The high prevalence of ST131 isolates resistant to fluoroquinolones and beta-lactams and their association with the carriage and dissemination of the extended-spectrum-beta-lactamase (ESBL) antibiotic resistance gene $bla_{CTX-M-15}$ have been documented (3–8). Based on their phylogenetic relationship, ST131 isolates have been divided into three main clades: clade A (predominantly represented by serotype O16:H5 *fimH41* allele isolates), clade B (predominantly represented by serotype O25b:H4 *fimH22* isolates), and clade C (predominantly represented

TABLE 1 Antibiotic susceptibility profile of *E. coli* ST131 isolates recovered in two hospitals in Armenia[a]

ID	Sequence type	Date[b]	Hospital	Serotype	*fimH*	AMP	PTZ	AMC	CAZ	CPM	Nor	LVX	AMK	IPM	MEM	CHL
ARM32	ST131	07/03/2019	H4	O16:H5	41	R	I	R	R	R	R	R	I	S	S	S
ARM42	ST131	15/04/2019	H5	O25b:H4	30	R	S	R	R	R	R	R	S	S	S	S
ARM46	ST131	11/03/2019	H5	O25b:H4	30	R	R	R	R	R	R	R	I	I	S	R
ARM86	ST131	09/01/2019	H4	O16:H5	41	R	R	R	R	R	R	R	I	S	S	S

[a]R, resistant; I, intermediate; S, sensitive; AMP, ampicillin (10 mg); PTZ, piperacillin-tazobactam (30/6 mg); AMC, amoxicillin (20 mg)-clavulanic acid (10 mg); CAZ, ceftazidime (10 mg); CPM, cefepime (30 mg); NOR, norfloxacin (10 mg); LVX, levofloxacin (5 mg); AMK, amikacin (30 mg); IMP, imipenem (10 mg); MEM, meropenem (10 mg); CHL, chloramphenicol (30 mg).
[b]Dates are given in the format day/month/year.

by serotype O25b:H4 *fimH30* isolates) (9). Furthermore, clade C is further divided into three subclades: C0 (*H30* lineage, which includes fluoroquinolone-sensitive isolates), C1 (*H30R* lineage, including fluoroquinolone-resistant isolates), and C2 (*H30Rx* lineage, including fluoroquinolone-resistant isolates with a G723A single nucleotide polymorphism [SNP] in the putative allantoin permease gene *ybbW* and carrying the $bla_{CTX-M-15}$ ESBL gene) (10–12). The carriage of the bla_{CTX-M} ESBL gene in ST131 isolates is linked to a specific incompatibility group F (IncF) conjugational plasmid type that is also associated with a particular clade C lineage. The most common of these plasmid types is Inc[F1:A2:B20], which is associated with the ST131 C1 lineage and carriage of the $bla_{CTX-M-27}$ gene, as well as Inc[F2:A1:B−], which is associated with the ST131 C2 lineage and carriage of the $bla_{CTX-M-15}$ gene (9, 12, 13).

Studies reporting on the transmission and the genetic makeup of *E. coli* ST131 are well documented globally, including those that have used a combination of standard molecular techniques such as PCR and whole-genome sequencing analysis using short-read sequencing platforms with high throughput and accurate base calling to infer bacterial genotype and phenotype (11, 14–16). Furthermore, advances in long-read sequencing have enabled us to resolve repetitive elements so that we can identify the structure and positioning of mobile genetic elements (MGE) in chromosomes and plasmids, such as antibiotic resistance and virulence genes (17–20). However, such genomic studies reporting whole-genome analysis of the *E. coli* ST131 lineage recovered in health care settings (including patients) in Armenia are absent. In this study, we report for the first time the whole-genome sequencing and analysis of the ST131 lineage of ESBL uropathogenic *E. coli* (UPEC) ST131 isolates recovered from patients' urine specimens in two hospitals in Armenia. To conduct the genomic analysis on Armenian ESBL-producing UPEC ST131, we used both Illumina and Oxford Nanopore Technologies (ONT) sequencing technologies to provide an in-depth genomic analysis identifying novel or rare genetic features within the Armenian isolates.

RESULTS

Isolates and antibiotic susceptibility testing. Four of 12 ESBL-producing UPEC isolates were identified as ST131 isolates (see Table S1 in the supplemental material). The four isolates identified as ST131 were received from medical microbiology laboratories (they were recovered from urine specimens from patients) of two hospitals, designated H4 (*n* = 2) and H5 (*n* = 2), in Yerevan, Armenia, between January 2019 and April 2019 (Table 1). These isolates were typed as O16:H5-*fimH41* (ARM32 and ARM86) and O25b:H4-*fimH30* (ARM42 and ARM46). All four Armenian *E. coli* ST131 isolates were resistant to the *β*-lactam antibiotics ampicillin and amoxicillin-clavulanic acid, the cephalosporin antibiotics ceftazidime and cefepime, and the fluoroquinolone antibiotics norfloxacin and levofloxacin. In addition, two isolates (ARM86 and ARM34) were resistant to the *β*-lactam antibiotic piperacillin-tazobactam; one isolate (ARM32) possessed intermediate resistance to piperacillin-tazobactam; one isolate (ARM46) was resistant to chloramphenicol; three isolates (ARM32, ARM46, and ARM86) possessed intermediate resistance to the aminoglycoside antibiotic amikacin; and one isolate (ARM46) possessed intermediate resistance to the carbapenem antibiotic imipenem. However, all Armenian *E. coli* ST131 isolates were sensitive to meropenem (a carbapenem antibiotic).

Phylogenetic analysis of ST131. To compare the Armenian *E. coli* ST131 isolates with those previously reported, a core SNP maximum-likelihood (ML) phylogenetic tree from 11,386 SNP sites with 99 recombination sites filtered out was constructed by aligning the short-read assemblies of our isolates to 2,496 *E. coli* ST131 isolate draft genomes retrieved

from the European Nucleotide Archive (ENA) database that had been recovered from 28 countries (Fig. 1). These isolates were recovered from seven different sources (agriculture animals, avians, domestic animals, the environment, humans, meat, and wild animals), identified as belonging to five clades, 13 serotypes, and 27 *fimH* types (Table 2). Two Armenian *fimH30* isolates (ARM42 and ARM46) were found within clade C2 and were phylogenetically closely related to human isolates recovered from the United States. Specifically, ARM42 was related to an isolate recovered from a blood sample in the United States (DABYPR01), and ARM46 was related to two isolates recovered from human stool samples in New York (DABHLM01 and DABHLJ01), as well as two isolates recovered from humans in New Zealand (ARM42 was related to DABNRT01, and ARM46 was related to DABNRO01, an isolate recovered from a blood sample). All isolates harbored the $bla_{CTX-M-15}$ gene. The Armenian *fimH41* isolates (ARM32 and ARM86) were found to belong to clade A and were phylogenetically closely related to human isolates from the United States that were recovered from urine samples in Pittsburgh, PA (DABAPB01), and the Northeast (DABYPM01), as well as human isolates from Australia that were recovered from fecal samples (DADOMG01 and DABGRC01) and a sample of lung fluid (DADPDJ01). Temporal analysis of the *E. coli* ST131 phylogenetic tree showed no temporal signal within the data set (Fig. S1), with a correlation coefficient of 0.2.

Accessory genome analysis revealed that 46,216 of 49,081 genes shared between 2,500 *E. coli* ST131 isolates were accessory genes. The t-SNE (*t*-distributed stochastic neighbor embedding) plot revealed that the majority of *E. coli* ST131 isolates shared many accessory genes with isolates from the same clade, indicating that accessory genes are clonally linked (Fig. 2). We found that isolates belonging to Armenian clade A (ARM32 and ARM86) shared many accessory genes ($r = 0.93$, Z score $= 1.94$, $P = 0.02$) compared to the Armenian clade C2 isolates (ARM42 and ARM46) ($r = 0.88$, Z score $= 0.97$, $P = 0.16$) (Table S4). ARM42 had a significantly high correlation in its accessory genome ($r > 0.93$, Z score $= 1.94$, $P < 0.02$) with an isolate recovered from a blood sample in Canada (DABRYR01) and an isolate recovered from a human in New Zealand (DABNRT01).

Virulence gene comparisons. Overall, we identified 191 virulence genes within 2,500 *E. coli* ST131 isolates, with the mean number of virulence genes being 87 (range, 50 to 117) (Table S5). Among the Armenian isolates, we found that those (ARM32 and ARM86) belonging to clade A had 104 and 105 virulence genes, respectively, which was significantly higher ($P < 0.01$) than the mean number of virulence genes in clade A (mean, 87; range, 60 to 105). Furthermore, among the Armenian isolates, we also found that two isolates (ARM42 and ARM46) belonging to clade C2 had 88 and 79 virulence genes, respectively, similar to the mean number of virulence genes in clade C2 (mean, 87; range, 51 to 112).

By constructing the virulence-associated-gene hierarchy cluster heat map, we determined that the Armenian isolates (ARM32 and ARM86) found in clade A clustered with clade C1 isolates DADOVK01 (Jaccard similarity index $= 0.87$), which was recovered from a rectal swab in Australia, and DABIKV01 (Jaccard similarity index $= 0.84$) and DABIJI01 (Jaccard similarity index $= 0.86$), which were recovered from a urine sample in Thailand (Fig. 3A). We found that the microcin H47 operon genes *mchB*, *mchC*, *mchF*, and *mcmA*, the S-fimbrial operon genes *sfaB*, *sfaC*, *sfaD*, *sfaE*, *sfaF*, *sfaG*, *sfaH*, and *sfaY*, the salmochelin operon, *iroBCDE*, and *iroN* were present in clade A Armenian isolates but were absent in other clade A isolates ($n = 375$) in the ENA database. Furthermore, the microcin H47 operon genes were identified in only one other ST131 isolate (DADZMQ01), recovered from Denmark. The S-fimbrial operon and salmochelin operon were present in only 0.2% and 4.3% of the ENA ST131 isolates, respectively. Within clade C2, ARM42 clustered with isolates designated DABNBN01 (Jaccard similarity index $= 0.96$) and DABXUG01 (Jaccard similarity index $= 0.95$) (recovered from urine in the United States) (Fig. 3B), whereas ARM46 clustered with a clade A isolate, DADRBK01 (Jaccard similarity index $= 0.89$) (recovered from urine in New Zealand) (Fig. 3C).

Antibiotic resistance genotype comparison. Among 2,500 ST131 isolates, we identified (in total) 350 antibiotic resistance genotypes (genes and mutations), with the mean number of antibiotic resistance genotypes being 65 (range, 52 to 79). Of these, 17 antibiotic resistance genotypes were determined to be a part of the ST131 core resistome (found in 100% of isolates), with a further 57 being found in more than 90% of the isolates (Table S6). The
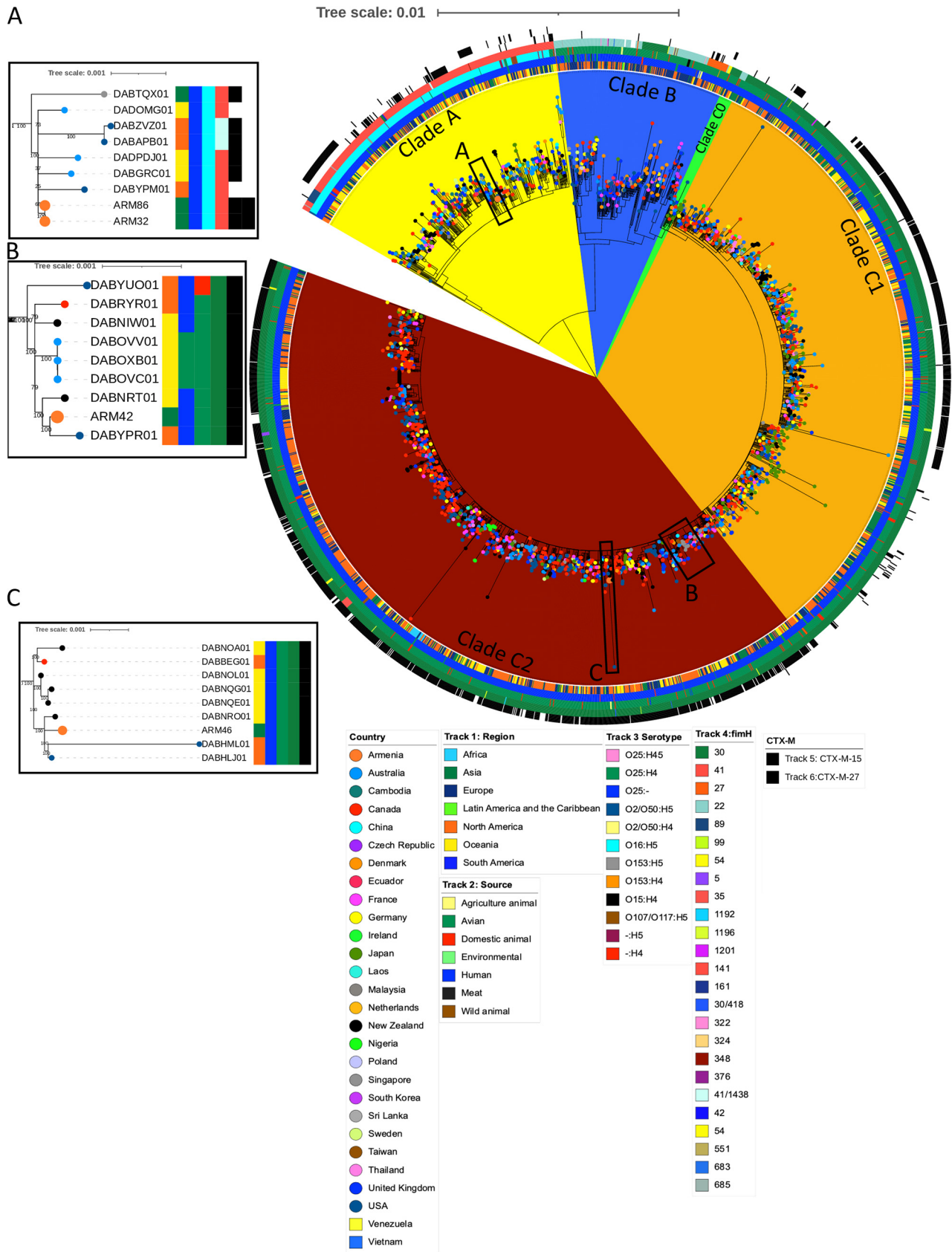
FIG 1 Maximum-likelihood phylogenetic tree of *E. coli* ST131 showing genetic diversity of the Armenian isolates.

TABLE 2 Lists of countries, sources, BAP clusters, and serotypes of ST131 isolates

Country (*n*)	Source (*n*)	Clade (*n*)	Serotype (*n*)	*fimH* allele (*n*)
Armenia (4)	Agriculture animal (3)	A (378)	−:H4 (97)	22 (148)
Australia (294)	Avian (54)	B (221)	−:H5 (9)	27 (26)
Cambodia (3)	Domestic animal (21)	C0 (17)	O107/O117:H5 (5)	30 (1,893)
Canada (581)	Environmental (2)	C1 (826)	O15:H4 (3)	35 (10)
China (10)	Human (2,413)	C2 (1,059)	O153:H4 (3)	41 (337)
Czech Republic (1)	Meat (2)		O153:H5 (3)	42 (1)
Denmark (30)	Wild animal (5)		O16:H5 (346)	54 (10)
Ecuador (10)			O2/O50:H4 (4)	89 (16)
France (118)			O2/O50:H5 (2)	99 (9)
Germany (123)			O25:- (8)	141 (6)
Ireland (1)			O25b:H4 (2,019)	161 (1)
Japan (195)			O25b:H45 (1)	234 (2)
Laos (6)				298 (1)
Malaysia (5)				322 (2)
Netherlands (16)				324 (1)
New Zealand (226)				348 (1)
Nigeria (23)				376 (1)
Poland (1)				551 (1)
Singapore (63)				683 (1)
South Korea (2)				685 (1)
Spain (52)				1192 (1)
Sri Lanka (8)				1196 (1)
Sweden (16)				1201 (2)
Taiwan (4)				22/685 (1)
Thailand (38)				30/418 (3)
United Kingdom (263)				41 (9)
USA (396)				41/1438 (2)
Venezuela (1)				
Vietnam (10)				

Armenian isolates of clade A (ARM32 and ARM86) possessed 69 and 73 antibiotic resistance genotypes, respectively, which are higher but not significantly higher ($P = 0.18$) than the mean number of antibiotic resistance genotypes in other clade A isolates (mean, 64; range, 55 to 73). The Armenian isolates of clade C2 (ARM42 and ARM46) possessed 69 and 63 antibiotic resistance genotypes, which were similar to the mean number of antibiotic resistance genotypes in clade C2 isolates (mean, 66; range, 56 to 79). Antibiotic resistance genotypes, *dfrA17*, *aadA5*, *qacEΔ1*, *sul1*, *bla*$_{CTX-M-15}$, and *Escherichia coli parC* S80I were found to be ubiquitous in Armenian isolates and were not part of the ST131 core resistome (<90% of the
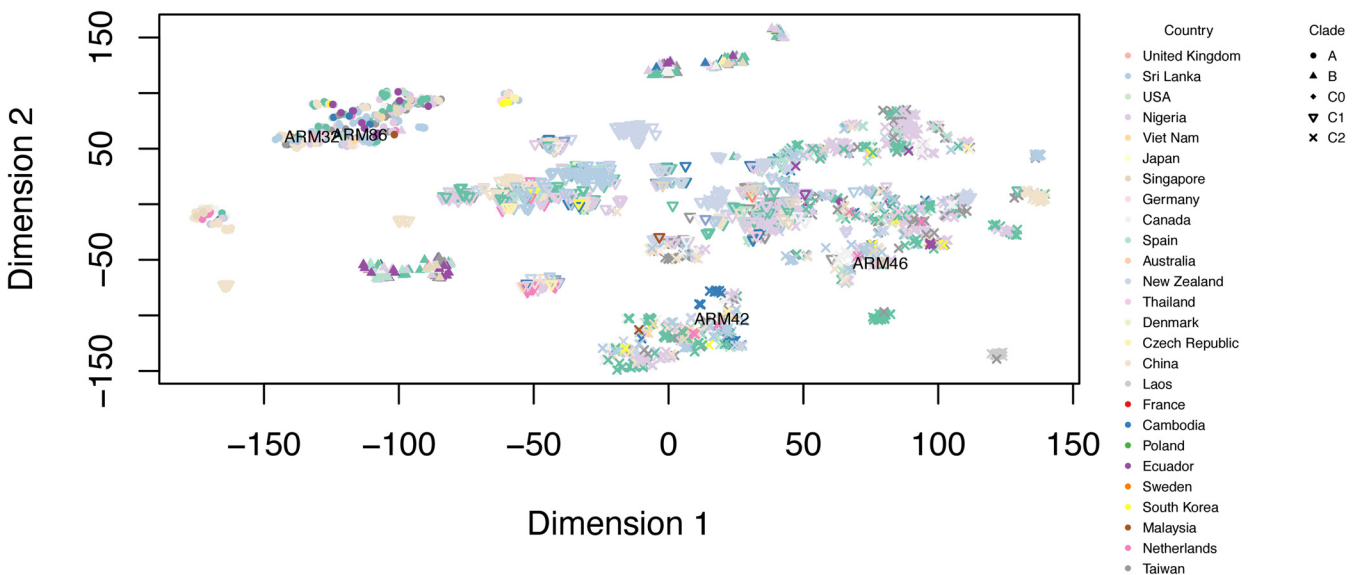
FIG 2 t-SNE plot of accessory genomes of ST131 isolates showing accessory genome differences.
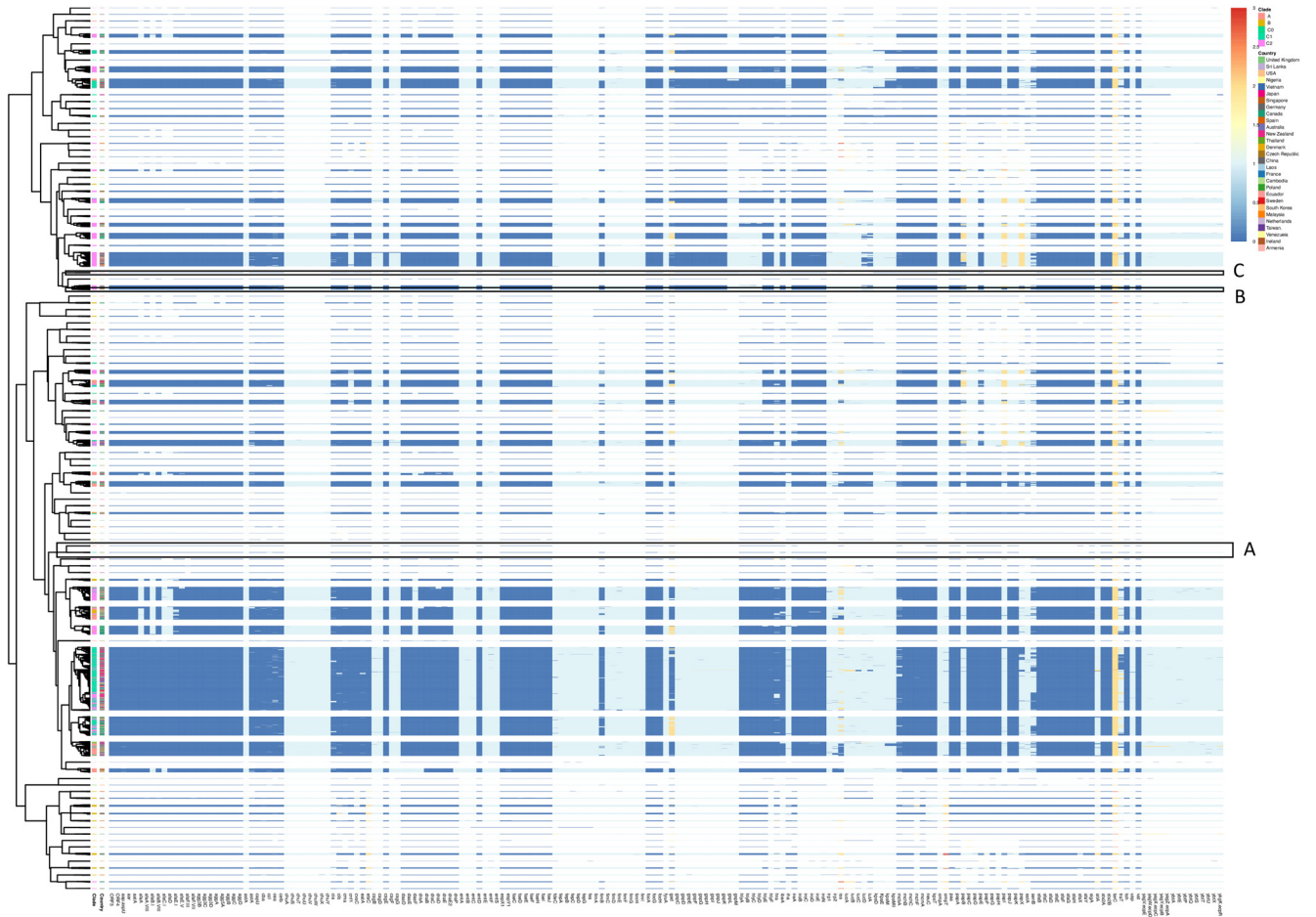
Total height of dendrogram: 0.76

A

Cluster height:0.22

B

Cluster height:0.07

C

Cluster height:0.14
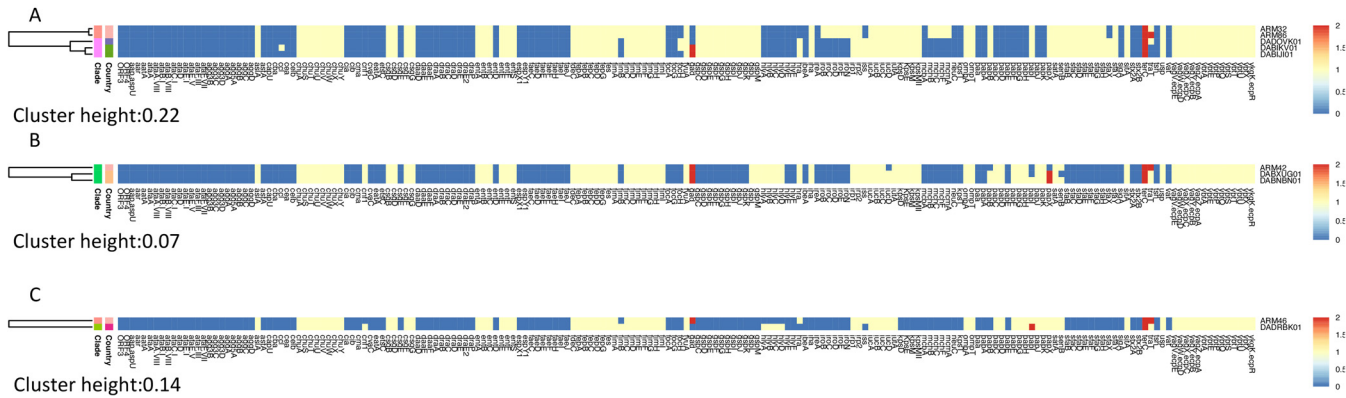
FIG 3 Hierarchy cluster heat map of *E. coli* ST131 isolates' virulence-associated gene profiles. (Top) Red, 3 copies of the gene; yellow, 2 copies of the gene; sky blue, 1 copy of the gene; dark blue, absence of gene. (A, B, and C) Red, 2 copies of the gene; yellow, 1 copy of gene; blue, absence of gene.

isolates) but were still commonly found in ST131 isolates (>40% of the isolates). ARM32, ARM42, and ARM86 also harbored bla_{OXA-1}, *aac(6′)-Ib-cr4*, *aac(3)-IIe*, *mphA*, and *catB3*. In addition, ARM42 and ARM86 harbored *tetA*; ARM32 and ARM86 harbored $bla_{CTX-M-27}$; and ARM86 harbored *APH(3″)-Ib*, *APH(6)-Id*, and *sul2*. The combination of $bla_{CTX-M-15}$ and $bla_{CTX-M-27}$ found in ARM32 and ARM86 was seen in only one other ST131 isolate (DADOWZ01), a clade C isolate recovered from a blood sample in Australia. The antibiotic resistance phenotypes of three Armenian isolates (ARM42, ARM32, and ARM86) correlated closely with the antibiotic resistance genotypes they harbored. However, we were unable to correlate ARM46 piperacillin-tazobactam, amikacin, imipenem, and chloramphenicol resistance phenotypes to a known antibiotic resistance genotype (Table 3).

TABLE 3 Antibiotic resistance genotypes of the Armenian *E. coli* ST131 isolates that were determined to match their phenotype[a]

| Isolate	Resistance phenotype	AMR genotype(s)
ARM32	AMP	OXA-1, CTX-M-27, CTX-M-15
	PTZ	OXA-1
	AMC	OXA-1, CTX-M-27, CTX-M-15
	CAZ	CTX-M-27, CTX-M-15
	CPM	OXA-1, CTX-M-27, CTX-M-15
	NOR	*parC* S80I, *gyrA* S83L, *gyrA* D87N
	LVX	*parC* S80I, *gyrA* S83L, *gyrA* D87N
	AMK	*aac(6')-lb-cr*
ARM42	AMP	CTX-M-15, OXA-1
	AMC	CTX-M-15, OXA-1
	CAZ	CTX-M-15
	CPM	CTX-M-15
	NOR	*parC* S80I, *gyrA* S83L, *gyrA* D87N
	LVX	*parC* S80I, *gyrA* S83L, *gyrA* D87N
ARM46	AMP	CTX-M-15
	PTZ	NA
	AMC	CTX-M-15
	CAZ	CTX-M-15
	CPM	CTX-M-15
	NOR	*parC* S80I, *gyrA* S83L, *gyrA* D87N
	LVX	*parC* S80I, *gyrA* S83L, *gyrA* D87N
	AMK	NA
	IPM	NA
	CHL	NA
ARM86	AMP	OXA-1, CTX-M-27, CTX-M-15
	PTZ	OXA-1
	AMC	OXA-1, CTX-M-27, CTX-M-15
	CAZ	CTX-M-27, CTX-M-15
	CPM	OXA-1, CTX-M-27, CTX-M-15
	NOR	*parC* S80I, *gyrA* S83L, *gyrA* D87N
	LVX	*parC* S80I, *gyrA* S83L, *gyrA* D87N
	AMK	*aac(6')-lb-cr*

[a]AMP, ampicillin; PTZ, piperacillin-tazobactam; AMC, amoxicillin-clavulanic acid; CAZ, ceftazidime; CPM, cefepime; NOR, norfloxacin; LVX, levofloxacin; AMK, amikacin; IMP, imipenem; MEM, meropenem; CHL, chloramphenicol; NA, not available.

We found that all Armenian isolates clustered based on having antibiotic resistance genotypes similar to those of ENA isolates found in clade C (Fig. 4). Based on their antibiotic resistance genotype, ARM32 and ARM42 (Jaccard similarity index = 0.97) clustered with DABADN01 (Jaccard similarity index = 0.96 and 0.99 for ARM32 and ARM42, respectively) and DAAZZW01 (Jaccard similarity index = 0.93 and 0.96) (recovered in Canada), BGWR01 and BGWS01 (both having a Jaccard similarity index of 0.96 and 0.99 for ARM32 and ARM42, respectively), BGWJ01 (Jaccard similarity index = 0.97), BGZC01, BGYZ01, and BGYQ01 (all having a Jaccard similarity index of 0.97 and 1 for ARM32 and ARM42, respectively) (recovered from Japan), and DABHAU01 (recovered in the United States) and DABFGX01 (recovered in the United Kingdom) (both having a Jaccard similarity index of 0.96 and 0.99 for ARM32 and ARM42, respectively) (Fig. 4A). ARM46 had the same antibiotic resistance genotype as isolates DABHAX01 and DABAPC01 (recovered in the United States) and DABGWM01 and DABGWL01 (recovered in Australia) (Jaccard similarity index = 1) (Fig. 4B). ARM46 also clustered with DAAZJF01 (Jaccard index = 0.98) (recovered in Germany), DABBEG01 (Jaccard similarity index =0.97) (Canada), and DABGJQ01 (Jaccard similarity index = 0.95) (France). Moreover, ARM86 clustered with isolate AAXCOJ01 (Jaccard similarity index = 0.97) and with isolates AAXCPJ01 (recovered in Nigeria) and AATDBO01 (recovered in the United States) (both having a Jaccard index of 0.99) (Fig. 4C).

Plasmid analysis of the Armenian ST131 isolates. Using hybrid genome assembly, we were able to obtain complete assemblies of the Armenian isolates' plasmids. Two out of
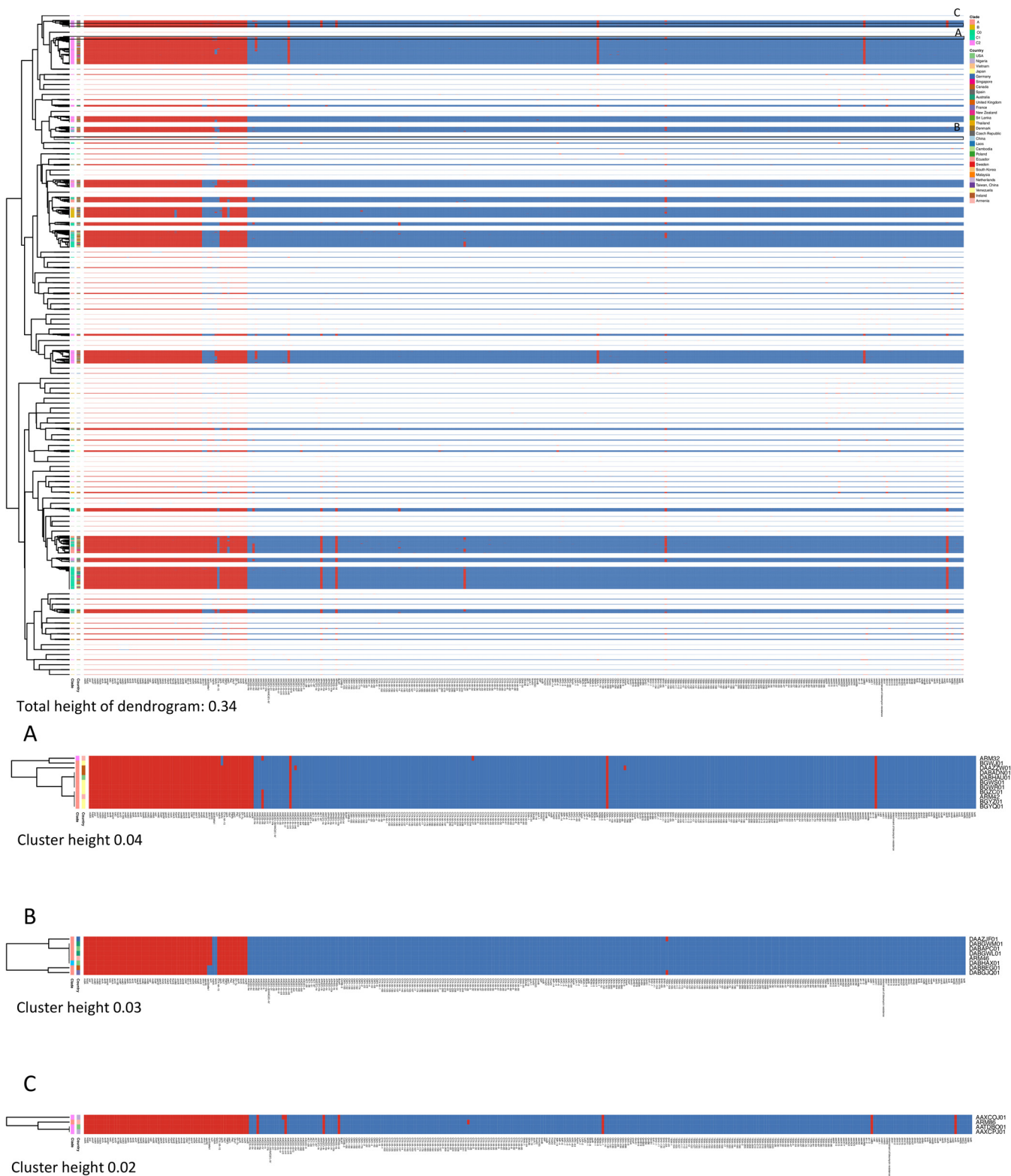
FIG 4 Hierarchy cluster heat map of *E. coli* ST131 isolates' antibiotic resistance genotype profiles. Red, presence of gene; blue, absence of gene.

four of the Armenian isolates (ARM42 and ARM86) harbored two plasmids, whereas ARM46 harbored three plasmids and ARM32 possessed only a single plasmid. Among the Armenian clade A isolates, the ARM32 plasmid (ARM32p) was typed as an IncF[F2:A2:B20] plasmid (Fig. 5A), ARM86 plasmid 1 (ARM86p1) was typed as an IncF[F1:A2:B20] plasmid (Fig. 5B), and ARM86 plasmid 2 (ARM86p2) was typed as an IncF[F2:A−:B−] plasmid (Fig. 5C). Among

8
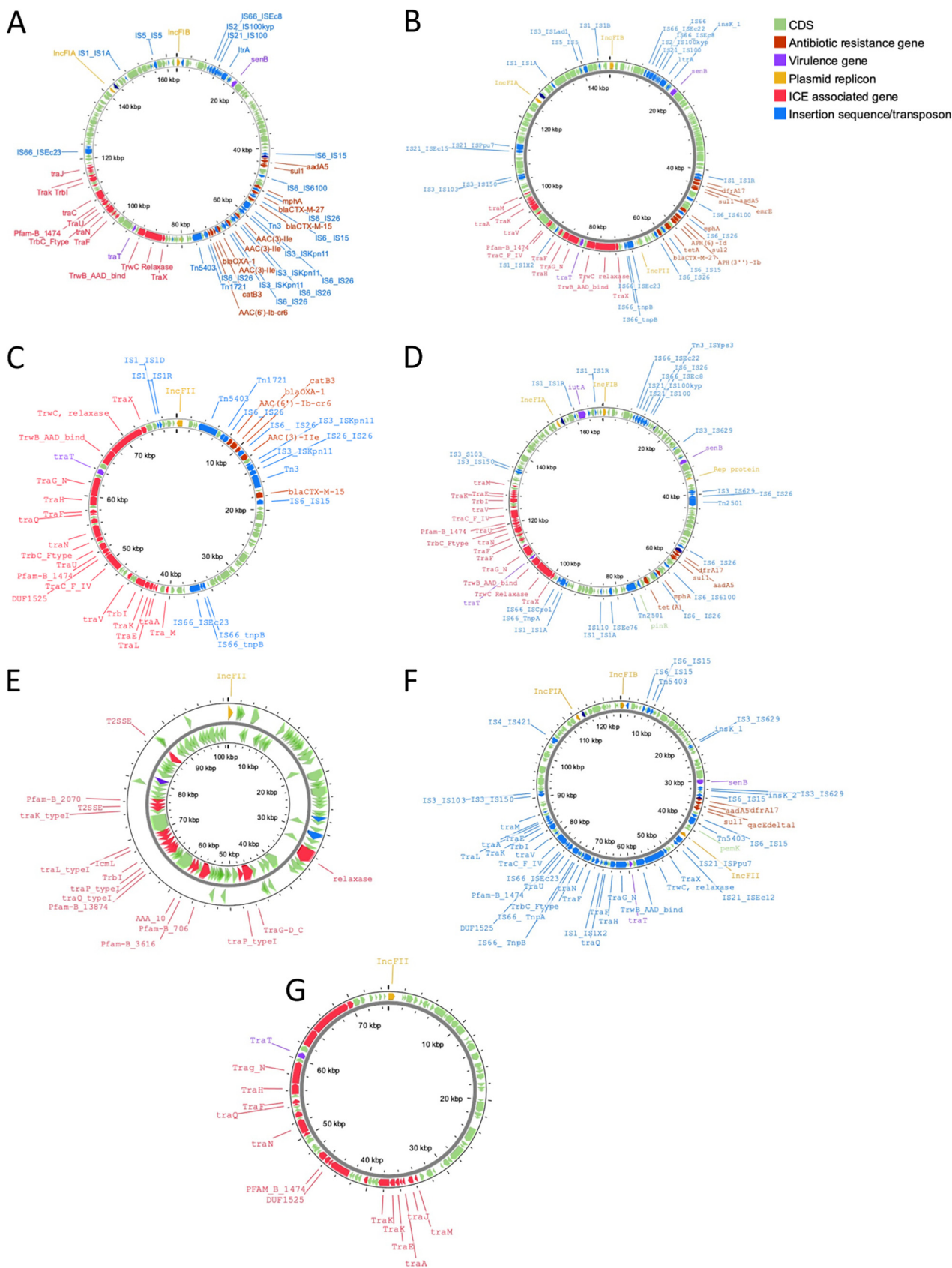
FIG 5 Armenian *E. coli* ST131 plasmids. (A) ARM32p1 IncF[F2:A2:B20]; (B) ARM86p1 IncF[F1:A2:B20]; (C) ARM86p2 IncF[F2:A−:B−]; (D) ARM42p1 IncF[F36/31:A4:B1]; (E) ARM42p2 IncC/B/O/K/Z; (F) ARM46p1 IncF[F1:A1:B16]; (G) ARM46p1[F2:A−:B−].

ARM42_IS6-IS26_chromosome

ARM86p2_IS6-IS26_plasmid

Gene feature

■ Antibiotic resistance gene

■ Insertion sequence
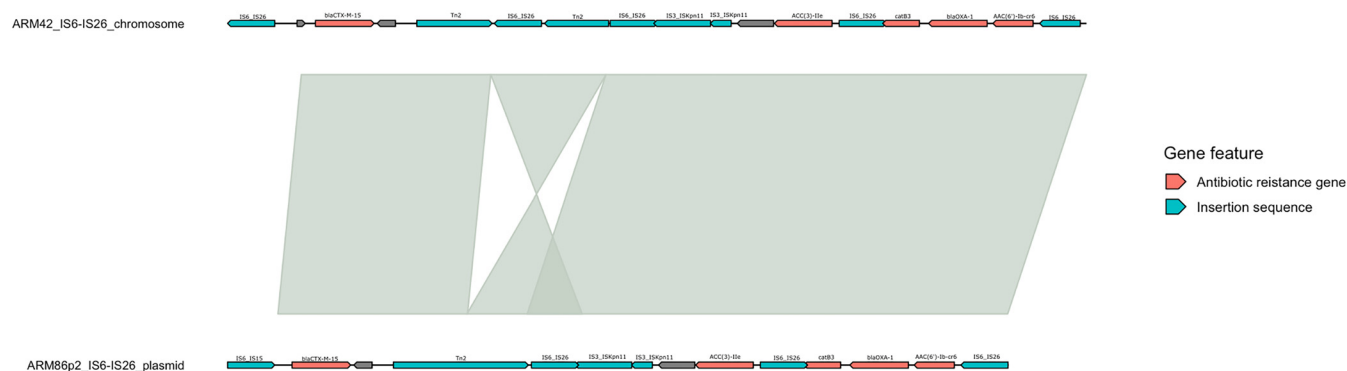
FIG 6 *E. coli* CTX-M-15 IS26 antibiotic resistance islands in ARM42 and ARM86 showing high genetic similarities.

the Armenian clade C2 isolates, ARM42 plasmid 1 (ARM42p1) was identified as an IncF[F31/36: A4:B1] plasmid (Fig. 5D), ARM42 plasmid 2 (ARM42p2) as an IncB/O/K/Z plasmid (Fig. 5E), ARM46 plasmid 1 (ARM46p1) as an IncF[F1:A1:B16] plasmid (Fig. 5F), ARM46 plasmid 2 (ARM46p2) as an IncF[F2:A-:B-] plasmid (Fig. 5G), and ARM46 plasmid 3 as a Col(MG828) plasmid which possessed only a plasmid replicon and a single hypothetical gene.

Five of the plasmids harbored antibiotic resistance genes within an IS26 genomic island, with ARM32p ($n = 13$) harboring the most antibiotic resistance genes, followed by ARM86p1 ($n = 10$), ARM86p1 ($n = 10$), ARM86p2 ($n = 5$), and ARM46p1 ($n = 4$). The ESBL gene $bla_{CTX-M-15}$ was present in the Armenian clade A isolate plasmids ARM32p and ARM86p2 but was chromosomal in clade C2 Armenian isolates. The $bla_{CTX-M-15}$ gene in ARM46 was integrated into the chromosome via an ISEcp1-CTX-M-15 transposition unit, whereas in ARM42, the $bla_{CTX-M-15}$ gene was integrated within the chromosome as part of the IS26 genomic island containing *aac(3)-IIe*, *catB3*, bla_{OXA-1}, and *aac(6')-Ib-cr6* antibiotic resistance genes. Furthermore, the $bla_{CTX-M-15}$ IS26 genomic island in ARM42 had a high degree of genetic synteny with the $bla_{CTX-M-15}$-IS26 genomic island found in ARM86p2 (Fig. 6) and high genetic similarities within the chromosome in complete *E. coli* ST131 genomes obtained from the National Center for Biotechnology Information (NCBI) nucleotide repository (Table S7).

Synteny analysis of the Armenian isolates' plasmids revealed that ARM32p had high genetic similarities with both ARM86p1 and ARM86p2 (Fig. 7). Further analysis of ARM32p identified two regions of reverse synteny with ARM86p2 adjacent to each other and flanked by genomic sequences that had synteny to ARM86p1. These regions of synteny consisted of ARM86p2 between its IncFII replicon and IS15 and contained the ARM86p2 antibiotic resistance genes $bla_{CTX-M-15}$, *aac(3)-IIe*, *cat3B*, bla_{OXA-1}, and *aac(6')-Ib-cr6*. The other region of synteny with ARM86p2 contained the integrating conjugative element (ICE)-associated genes. There were multiple regions of synteny to ARM86p1 within the ends and middle sections of the same ICE-associated gene region. There was also a change in sequence orientation of the ARM86p IS26-$bla_{CTX-M-27}$-IS903B-IS15 region in ARM32p. ARM32p had all the antibiotic genes found in ARM86p2 but had only 6 of the 10 antibiotic resistance genes found in ARM86p1; it was missing *tetA*, *APH(6)-Id*, *APH(3')-Ib*, and *sul2*. Moreover, ARM32p possessed 3 copies in direct repeat of a genetic region found in ARM86p1 consisting of IS26, ISKpn11, the AAA family ATPase gene, *aac(3)*, *IIe*, and IS26.

Our data analysis showed high genetic similarities of *E. coli* ST131 draft genomes obtained from the ENA database to the Armenian isolates' plasmids. We found that ARM42p1 had a high score (>0.999) of genetic identity to clade C2 ENA isolates that harbored the IncF [F31/36:A4:B1] genotype and had the same antimicrobial resistance (AMR) as ENA isolates recovered from Japan, Canada, the United Kingdom, New Zealand, Australia, and the United States. The Japanese isolate BGZA01 had the highest genetic similarity contained in its genome, whereas the Canadian isolate DABABP01 (recovered from a blood sample) had the highest BLAST coverage (0.93) relative to ARM42p1. ARM46p1 had a high score (>0.999) of genetic identity to clade C2 ENA isolates that possessed the IncF[F1:A1:B16] genotype and had the same AMR genes as ENA isolates recovered from France, Canada, the United Kingdom, New Zealand, Australia, and the United States. The U.S. isolate AATDBO01
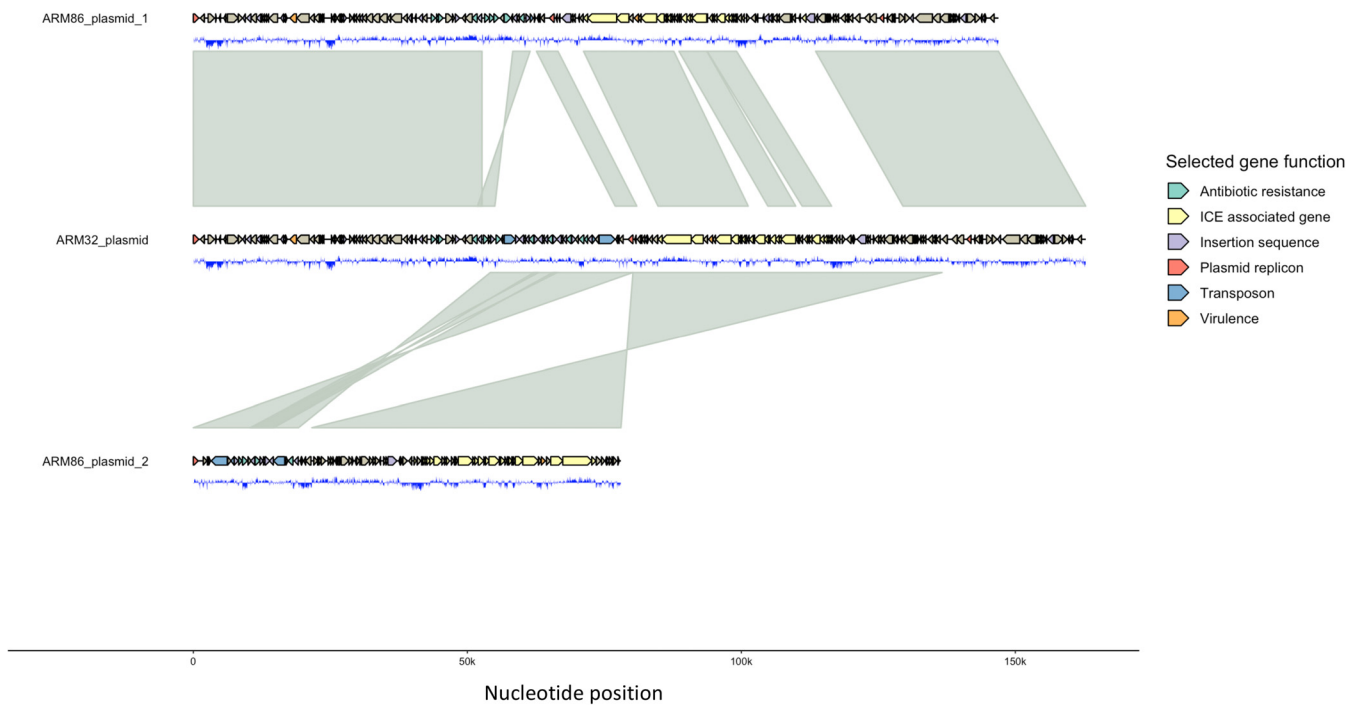
FIG 7 Plasmid symmetry comparison of ARM32 and ARM86 showing two separate recombination cointegration events from ARM86 plasmid 2 into the ARM86 plasmid backbone to form the ARM32 plasmid.

(recovered from a blood sample) had the highest genetic identity score and the Australian isolate DADPEZ01 had the highest BLAST coverage (0.92) relative to ARM46p1. For the other Armenian isolate plasmids, we did not find genetic identities higher than >0.999 contained within the ENA isolates. The ENA isolate DADOWZ01, which had both the $bla_{CTX-M-15}$ and $bla_{CTX-M-27}$ genes, had the same IncF type as ARM32 but was not genetically close to ARM32p (identity score, 0.983; BLAST coverage, 0.64), ARM86p1 (identity score, 0.989; BLAST coverage, 0.47), and ARM86p2 (identity score, 0.962; BLAST coverage, 0.75).

DISCUSSION

In this study, we report for the first time the antibiotic resistance profiles and genetic features of four UPEC ST131 isolates recovered from patients in Armenia, where genomic surveillance is lacking. By combining the short-read sequencing for high base calling accuracy to infer Armenian isolates' genotype and phylogeny and long-read sequencing to resolve repetitive elements and identify structures and positioning of MGE within the chromosome and plasmid, we were able to perform an in-depth genomic analysis of the Armenian *E. coli* ST131 isolates (20). We found that two Armenian ST131 isolates were genotyped as *H*30Rx/clade C2 (ARM42 and ARM46), one of the most prevalent *E. coli* ST131 genotypes that cause UTI globally (2, 21). The other two isolates were genotyped as *H*41/clade A (ARM32 and ARM86), which has recently emerged as one of the most common ST131 genotypes causing UTI in Australia, New Zealand, and China (15, 22, 23). Intriguingly, ARM32 and ARM86 are phylogenetically closely related to each other and share many of the same accessory genes but were recovered from patients in two different hospitals, which may indicate a wide community transmission of this strain. Furthermore, we found that the Armenian isolates were phylogenetically more closely related to isolates recovered from Australia, New Zealand, and the United States and were more likely to share accessory genes such as those for virulence factors and antibiotic resistance genotypes than isolates recovered from other countries. Moreover, using long-read sequencing, we were able to determine that genetic elements in the ENA isolates (those that were recovered from countries such as Australia, New Zealand, and the United States) had high genetic similarities with the Armenian isolates. We hypothesize that because there are large Armenian diasporas in Australia, New Zealand, and especially

the United States, with close family links in Armenia, it seems plausible that the transmission could have occurred through members of these Armenian communities (24, 25).

The antibiotic resistance phenotypes of Armenian ST131 isolates were consistent with that of ST131 isolates except in the case of ARM46, which demonstrated intermediate resistance to the carbapenem antibiotic imipenem (26, 27). Antibiotic resistance genotyping also showed that carbapenem and chloramphenicol resistance was absent in ARM46, potentially indicating that it may have developed resistance via an unknown mechanism. Moreover, we were unable to determine the antibiotic resistance genotype for amikacin and piperacillin-tazobactam in ARM46, which, however, was determined for other Armenian isolates due to the identification of associated *aac(6′)-Ib-cr* and *bla*$_{OXA-1}$ genes. For all isolates, ESBL production was linked to the *bla*$_{CTX-M-15}$ gene, which was plasmid borne in Armenian *H*41 isolates and chromosomal in *H*30Rx isolates. Chromosomal *bla*$_{CTX-M-15}$ is generally considered rare in *E. coli* ST131 isolates, as IncF *bla*$_{CTX-M-15}$ plasmids are very stable within the population even in the absence of antibiotic pressure (28–30). However, the general prevalence of chromosomal *bla*$_{CTX-M-15}$ is still relatively unknown in whole-genome sequencing studies due to the use of short-read sequencing platforms, which have difficulties in differentiating between plasmid-borne and chromosomal genetic features in comparison to long-read sequencing (17, 31). The chromosomal integration of *bla*$_{CTX-M-15}$ in ARM46 was related to IS*Ecp1-bla*$_{CTX-M-15}$, which previously was identified as chromosomal in ST131 isolates recovered in Zambia and Ireland (30, 32), whereas the chromosomal integration of *bla*$_{CTX-M-15}$ in ARM42 was found within an IS*26* antibiotic resistance genomic island which was also present in ST131 isolates' chromosomes (obtained from the NCBI nucleotide database). Moreover, this IS*26* antibiotic resistance island was also present in the Armenian *H*41 isolates' plasmids, which may suggest that this resistance genomic island was common among the Armenian ST131 isolates.

Moreover, we were able to determine that the Armenian *H*41 isolates harbored both *bla*$_{CTX-M-15}$ and *bla*$_{CTX-M-27}$ genes. Intriguingly, we found that the ARM32 IncF[F2:A2:B20] plasmid harboring both *bla*$_{CTX-M-15}$ and *bla*$_{CTX-M-27}$ is a fusion of ARM86 *bla*$_{CTX-M-15}$ on the IncF[F2:A−:B−] plasmid and *bla*$_{CTX-M-27}$ on the IncF[F1:A2:B20] plasmid, most likely from a recent recombination event between the plasmids' IS*26* antibiotic genomic islands (33, 34). Recombination events between IncF plasmids have been reported; however, the movement of the *bla*$_{CTX-M-15}$ gene into the *bla*$_{CTX-M-27}$ plasmid backbone in the *E. coli* ST131 genetic background is undocumented (13, 35, 36). The fusion plasmid of ARM32 also harbored many of the antibiotic resistance genes on the two plasmids found in ARM86, which may make the horizontal transfer of multiple antibiotic resistance genes easier. Furthermore, we found an increased copy number, detected by long-read sequencing with the ability to resolve repetitive elements, of the aminoglycoside antibiotic resistance genes *aac(3)-IIe* (1 copy in ARM86 and 3 copies in ARM32), which has been shown to increase the MIC of tobramycin (37–39). The evolution observed between these ARM32 and ARM86 plasmids demonstrates that recombination of *bla*$_{CTX-M-15}$ can occur in the *bla*$_{CTX-M-27}$ plasmid backbone within ST131-*fimH*41 backgrounds, which has not been seen before in other ST131 lineages, as they are restricted to carrying certain IncF plasmid types (13). Only one ENA isolate (DADOWZ01) that harbored both *bla*$_{CTX-M-15}$ and *bla*$_{CTX-M-27}$ genes possessed the IncF[F2:A2:B20] type, which suggests that this plasmid genotype may be the only genotype that can harbor both *bla*$_{CTX-M-15}$ and *bla*$_{CTX-M-27}$. Moreover, as antibiotic resistance genes within IncF plasmids are acquired via homology recombination with common features in the IncF plasmid rather than via transposition, the development of the IncF[F2:A2:B20] genotype is most likely to occur due to allele exchange of two different IncF plasmids' IncFII replicons (34, 36, 40).

The Armenian *H*41 isolates harbored more virulence genes than ENA isolates that belonged to the same genotype. In particular, the Armenian isolates harbored a genomic island which contained the S-fimbrial operon (encoding adhesin factor), the salmochelin operon (encoding an iron siderophore), and a microcin H47 operon (which encodes a bacteriocin that inhibits other enteric pathogens) (41); this genomic island is known but uncommon in ST131 isolates. Previous reports of isolates harboring the microcin H47–salmochelin–S-fimbrial genomic island generally showed increased virulence, suggesting

that these genes provide a fitness advantage for *E. coli*, contributing to its survival and acquisition of virulence factors (42, 43). Moreover, salmochelin and microcin H47 genes may have a role in promoting urinary tract colonization, as observed by the upregulation of salmochelin and microcin H47 genes when isolates are grown in urine (42–45). As ST131 isolates are considered the most common sequence type causing UTI, the lack of this pathogenicity island suggests that these virulence factors may not be entirely necessary for them to be successful in causing UTI. Moreover, the mechanism of the success of ST131 in causing UTI compared to other *E. coli* sequence types is still relatively unknown (46–49).

In summary, we report for the first time genetic features and transmission of ESBL-producing *E. coli* ST131 isolates recovered in Armenia using both short- and long-read sequencing technologies to provide in-depth genomic analysis to better understand Armenian isolates' genotype, phylogeny, and MGE structure and position within the chromosome and plasmid. Although we had only 4 isolates belonging to two different ST131 genotypes, we were able to identify the close relatedness of these isolates to those recovered in the United State, Australia, and New Zealand. Our data show that Armenian isolates included in this study possessed genetic features not commonly identified in other *E. coli* ST131 isolates, including the $bla_{CTX-M-15}$ gene integrated into the ST131-*H*30Rx isolates' chromosomes and the acquisition of the pathogenicity island containing the S fimbrial, salmochelin siderophore, and microcin H47 virulence genes in ST131-*H*41 isolates. Moreover, using hybrid assembly of short- and long-read sequencing in ST131-*H*41 isolates, we detected the formation of IncF[F2:A2:B20] containing both $bla_{CTX-M-15}$ and $bla_{CTX-M-27}$ derived from the recombination of genes from an IncF[F2:A−:B−] $bla_{CTX-M-15}$-associated plasmid into an IncF[F1:A2:B20] $bla_{CTX-M-27}$-associated plasmid backbone. Further genomic surveillance of ESBL *E. coli* ST131 isolates in Armenia is necessary to determine whether the genetic features identified in our study are widely disseminated in other *E. coli* ST131 UTI isolates, which could contribute to global AMR transmission.

MATERIALS AND METHODS

Bacterial isolation and identification. Twelve ESBL-producing UPEC isolates were received from medical microbiology laboratories of five hospitals in Armenia between January 2019 and August 2019 (Table S1) (50). All isolates were recovered from urine specimens from hospitalized patients. The isolates were identified as *E. coli* using matrix-assisted laser desorption ionization–time-of-flight mass spectroscopy (MALDI-TOF MS) as described previously (51). Four of 12 isolates were identified as *E. coli* ST131 by multilocus sequence typing (MLST) from whole-genome sequencing using mlst (https://github.com/tseemann/mlst [accessed July 2021]) and were selected for use in this study.

Antibiotic susceptibility testing. The identification and susceptibility profiles of these four and other *E.coli* isolates were reported previously (50); in brief, each isolate was tested against a panel of 10 antibiotics using the disk diffusion method. The antibiotics included were ampicillin (10 mg), piperacillin-tazobactam (30/6 mg), amoxicillin-clavulanic acid (20/10 mg), ceftazidime (10 mg), cefepime (30 mg), norfloxacin (10 mg), levofloxacin (5 mg), amikacin (30 mg), imipenem (10 mg), meropenem (10 mg), and chloramphenicol (30 mg) (Mast Group, Merseyside, UK) according to the European Committee on Antimicrobial Susceptibility Testing (EUCAST) protocol (52). Isolates were determined to be ESBL producing if they showed resistance to ceftazidime and cefepime.

Genome sequencing and assembly. The 12 *E. coli* isolates were whole-genome sequenced using the Illumina HiSeq platform. Genomic DNA was extracted using the TIANamp bacterial DNA kit (Tiangen, China), and paired-end sequencing libraries were constructed using Nextera XT DNA sample preparation kits or TruSeq DNA HT sample preparation kits (Illumina, USA) following the manufacturer's instructions.

The quality of short reads was checked using FastQC, with low-quality reads trimmed by Trimmomatic (53). Trimmed reads were *de novo* assembled using SPAdes (54).

Four of 12 isolates that were identified as *E. coli* ST131 were further sequenced for detailed chromosomal and plasmid genetic investigation using the Oxford Nanopore Technologies MinIon 1B sequencing platform. Genomic DNA for long-read sequencing was extracted using a Qiagen MagAttract HMW DNA kit (Qiagen, Germany), and the long-read library was constructed using a ligation sequencing kit (SQK-LSK109) (Oxford Nanopore Technologies, UK) and run on an R9.4.1 flow cell (Oxford Nanopore Technologies, UK). Guppy v6.3.4 was used for long-read base calling, filtering out reads with a Q score of <9 (55).

Long and short reads were combined for hybrid genome assembly using two different assembly pipelines depending on the sequencing depth of long reads. All isolates were assembled using the Unicycler assembly pipeline for the detection of plasmids less than <10 kb (56, 57). The genome of one isolate (ARM42) was assembled using only Unicycler. Isolates ARM32, ARM46, and ARM86 (which had a long-read sequence depth greater than 50×) were also assembled using Canu, Flye, Miniasm+Minipolish, Raven, NECAT, and NextDenovo/NextPolish (58–63). A consensus of the long-read assemblies was constructed via Trycycler (64). The final Trycycler assemblies were polished with the short-read sequences of the same isolates using Polypolish and POLCA (65, 66).

Genome selection for phylogenetic and genomic comparison. A total of 2,496 draft *E. coli* ST131 genomes obtained from the ENA database were selected for genomic comparison (accessed January 2022) (Table S3). The isolates were selected based on the criteria that they (i) belonged to ST131 as determined with the mlst (https://github.com/tseemann/mlst accessed July 2021) typing scheme (PubMLST) (accessed January 2022) and (ii) were accompanied by data on their isolation, including date, source, and country.

Phylogenetic tree analysis. To reconstruct the core SNP maximum-likelihood (ML) phylogenetic tree, the short sequence reads of *E. coli* ST131 isolates recovered from Armenia and the isolates from the ENA database were first aligned with the ST131 reference genome EC958 (accession no. HG941718.1) using Snippy (https://github.com/tseemann/snippy). Recombination within the genome was filtered out using Gubbins (67). A phylogenetic tree was constructed using IQ-TREE v2.1.2 with the best model selected by ModelFinder and with ultrafast bootstrap replication set to 1,000 (68–70). The phylogenetic tree was visualized and annotated using iTOL (71). The temporal signal of the phylogenetic tree was determined by TempEst (72).

Genome annotation. *E. coli* ST131 isolate genomes were annotated using Prokka (73). The serotype was identified using ECtyper (74). Virulence-associated genes for each isolate were screened using ABRicate software (https://github.com/tseemann/ABRicate) in conjunction with the VirulenceFinder and VFDB database combined for virulence detection. Antibiotic resistance genotype was identified using Resistance Gene Identifier (RGI) software in conjugation with the Comprehensive Antibiotic Resistance Database (CARD) (accessed January 2022) (75). Antibiotic resistance related to deletion of outer membrane protein genes was also detected (76, 77). A hierarchy cluster heat map of isolated antibiotic resistance genotypes and virulence-associated genes was constructed using the R package Pheatmap. The *fimH* allele type, presence of the *ybbW* G723A SNP in *H*30Rx isolates, and IncF replicon sequence type (RST) were determined using SKA (https://github.com/simonrharris/SKA) in conjunction with the FimTyper and pMLST IncF RST database (accessed August 2022) (78–80).

For Armenian *E. coli* ST131 isolates, plasmids found in the hybrid genome assembly were typed using MOB-suite and pMLST (81, 82). ABRicate was used to screen for plasmid replicon and insertion sequences (IS) using the PlasmidFinder and ISfinder database (83, 84). ICE-associated genes were screened using ICEfinder (85). Plasmid structure was visualized using Proksee (https://proksee.ca/).

Pangenome analysis. The *E. coli* ST131 pangenome was constructed using both Armenian and ENA isolates combined using Roary with the BLASTp percentage identity cutoff set at 90% (86). A t-SNE plot of the accessory gene was plotted using the R package Rtsne (https://github.com/jkrijthe/Rtsne) and the pairwise Pearson correlation using the R package Hmisc (https://cran.r-project.org/web/packages/Hmisc/index.html).

Genetic synteny and plasmid similarity analyses. Plasmid synteny between Armenian isolates was identified from the hybrid genome assembly using minimap2 and visualized using the R package gggenomes (https://github.com/thackl/gggenomes) (87). Mash Screen was used to determine whether the *E. coli* ST131 isolates' draft genomes contained sequences matching those in the hybrid assembly of the Armenian isolates' plasmids, with only those with a Mash Screen identity of >0.995 reported (88). Pyani was used to measure BLAST coverage of the sequence in the ENA database to the Armenian isolates' plasmids (89).

Statistical analysis. Student's *t* test was used to determine the significant difference between the number of virulence and antibiotic resistance genotypes between isolates. Z-scored normalization was used to determine the significance between accessory gene correlation data. A Jaccard similarity index was used to determine the similarity coefficient between two isolates' virulence genes and antibiotic resistance gene profiles.

Data availability. The short- and long-read sequencing data generated in this study were deposited in the ENA database under the accession numbers ERS14467679 (ARM32), ERS14467680 (ARM42), ERS14467681 (ARM46), and ERS14467682 (ARM86). Individual accession numbers for the genome sequencing data are included in Table S2. Genome assemblies of the isolates are included in Table S2 and deposited in the ENA database under the accession numbers GCA_951802855 (ARM32), GCA_955652485 (ARM42), GCA_951803545 (ARM46), and GCA_951802865 (ARM86). Both Illumina short reads and Oxford Nanopore long reads were uploaded to the ENA repository (project PRJEB51925) (Table S2).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 0.6 MB.

SUPPLEMENTAL FILE 2, XLS file, 7 MB.

SUPPLEMENTAL FILE 3, XLSX file, 5.9 MB.

ACKNOWLEDGMENT

REFERENCES

1. Pitout JDD, DeVinney R. 2017. Escherichia coli ST131: a multidrug-resistant clone primed for global domination. F1000Res 6:195. https://doi.org/10.12688/f1000research.10609.1.

2. Nicolas-Chanoine M-H, Bertrand X, Madec J-Y. 2014. Escherichia coli ST131, an intriguing clonal group. Clin Microbiol Rev 27:543–574. https://doi.org/10.1128/CMR.00125-13.

3. Liss MA, Peterson EM, Johnston B, Osann K, Johnson JR. 2013. Prevalence of ST131 among fluoroquinolone-resistant Escherichia coli obtained from rectal swabs before transrectal prostate biopsy. Urology 81:548–556. https://doi.org/10.1016/j.urology.2012.10.056.

4. Rogers BA, Sidjabat HE, Paterson DL. 2011. Escherichia coli O25b-ST131: a pandemic, multiresistant, community-associated strain. J Antimicrob Chemother 66:1–14. https://doi.org/10.1093/jac/dkq415.

5. Wang J-L, Ko W-C, Hung C-H, Cheng M-F, Wang H-Y, Shiau Y-R, Lai J-F, Huang I-W, Hsieh L-Y, Lauderdale T-LY, Tsar Hospitals. 2021. Temporal trend of ST131 clone among urinary Escherichia coli isolates in the community: a Taiwan national surveillance from 2002 to 2016. Microorganisms 9:963. https://doi.org/10.3390/microorganisms9050963.

6. Castanheira M, Simner PJ, Bradford PA. 2021. Extended-spectrum β-lactamases: an update on their characteristics, epidemiology and detection. JAC Antimicrob Resist 3:dlab092. https://doi.org/10.1093/jacamr/dlab092.

7. Mathers AJ, Peirano G, Pitout JDD. 2015. The role of epidemic resistance plasmids and international high-risk clones in the spread of multidrug-resistant Enterobacteriaceae. Clin Microbiol Rev 28:565–591. https://doi.org/10.1128/CMR.00116-14.

8. Merino I, Hernández-García M, Turrientes M-C, Pérez-Viso B, López-Fresneña N, Diaz-Agero C, Maechler F, Fankhauser-Rodriguez C, Kola A, Schrenzel J, Harbarth S, Bonten M, Gastmeier P, Canton R, Ruiz-Garbajosa P, Desilets M, Dul S, Scherrer-Muller F, Huttner B, Uçkay I, Prendki V, Renzi G, R-GNOSIS Study Group. 2018. Emergence of ESBL-producing Escherichia coli ST131-C1-M27 clade colonizing patients in Europe. J Antimicrob Chemother 73: 2973–2980. https://doi.org/10.1093/jac/dky296.

9. Stoesser N, Sheppard AE, Pankhurst L, De Maio N, Moore CE, Sebra R, Turner P, Anson LW, Kasarskis A, Batty EM, Kos V, Wilson DJ, Phetsouvanh R, Wyllie D, Sokurenko E, Manges AR, Johnson TJ, Price LB, Peto TEA, Johnson JR, Didelot X, Walker AS, Crook DW, Modernizing Medical Microbiology Informatics Group (MMMIG). 2016. Evolutionary history of the global emergence of the Escherichia coli epidemic clone ST131. mBio 7:e02162-15. https://doi.org/10.1128/mBio.02162-15.

10. Price LB, Johnson JR, Aziz M, Clabots C, Johnston B, Tchesnokova V, Nordstrom L, Billig M, Chattopadhyay S, Stegger M, Andersen PS, Pearson T, Riddell K, Rogers P, Scholes D, Kahl B, Keim P, Sokurenko EV. 2013. The epidemic of extended-spectrum-β-lactamase-producing Escherichia coli ST131 is driven by a single highly pathogenic subclone, H30-Rx. mBio 4: e00377-13. https://doi.org/10.1128/mBio.00377-13.

11. Petty NK, Ben Zakour NL, Stanton-Cook M, Skippington E, Totsika M, Forde BM, Phan M-D, Gomes Moriel D, Peters KM, Davies M, Rogers BA, Dougan G, Rodriguez-Baño J, Pascual A, Pitout JDD, Upton M, Paterson DL, Walsh TR, Schembri MA, Beatson SA. 2014. Global dissemination of a multidrug resistant Escherichia coli clone. Proc Natl Acad Sci U S A 111: 5694–5699. https://doi.org/10.1073/pnas.1322678111.

12. Johnson TJ, Danzeisen JL, Youmans B, Case K, Llop K, Munoz-Aguayo J, Flores-Figueroa C, Aziz M, Stoesser N, Sokurenko E, Price LB, Johnson JR. 2016. Separate F-type plasmids have shaped the evolution of the H30 subclone of Escherichia coli sequence type 131. mSphere 1:e00121-16. https://doi.org/10.1128/mSphere.00121-16.

13. Kondratyeva K, Salmon-Divon M, Navon-Venezia S. 2020. Meta-analysis of pandemic Escherichia coli ST131 plasmidome proves restricted plasmid-clade associations. Sci Rep 10:36. https://doi.org/10.1038/s41598-019-56763-7.

14. Alqasim A, Abu Jaffal A, Alyousef AA. 2020. Prevalence and molecular characteristics of sequence type 131 clone among clinical uropathogenic Escherichia coli isolates in Riyadh, Saudi Arabia. Saudi J Biol Sci 27:296–302. https://doi.org/10.1016/j.sjbs.2019.09.020.

15. Li D, Wyrsch ER, Elankumaran P, Dolejska M, Marenda MS, Browning GF, Bushell RN, McKinnon J, Chowdhury PR, Hitchick N, Miller N, Donner E, Drigo B, Baker D, Charles IG, Kudinha T, Jarocki VM, Djordjevic SP. 2021. Genomic comparisons of Escherichia coli ST131 from Australia. Microb Genom 7:e000721. https://doi.org/10.1099/mgen.0.000721.

16. Chen SL, Ding Y, Apisarnthanarak A, Kalimuddin S, Archuleta S, Omar SFS, De PP, Koh TH, Chew KL, Atiya N, Suwantarat N, Velayuthan RD, Wong JGX, Lye DC. 2019. 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 9:13245. https://doi.org/10.1038/s41598-019-49467-5.

17. Decano AG, Ludden C, Feltwell T, Judge K, Parkhill J, Downing T. 2019. Complete assembly of Escherichia coli sequence type 131 genomes using long reads demonstrates antibiotic resistance gene variation within diverse plasmid and chromosomal contexts. mSphere 4:e00130-19. https://doi.org/10.1128/mSphere.00130-19.

18. da Cruz Campos AC, Cavallo FM, Andrade NL, van Dijl JM, Couto N, Zrimec J, Lo Ten Foe JR, Rosa ACP, Damasco PV, Friedrich AW, Chlebowicz-Flissikowska MA, Rossen JWA. 2019. Determining the virulence properties of Escherichia coli ST131 containing bacteriocin-encoding plasmids using short- and long-read sequencing and comparing them with those of other E. coli lineages. Microorganisms 7:534. https://doi.org/10.3390/microorganisms7110534.

19. Decano AG, Downing T. 2019. An Escherichia coli ST131 pangenome atlas reveals population structure and evolution across 4,071 isolates. Sci Rep 9:17394. https://doi.org/10.1038/s41598-019-54004-5.

20. Huisman JS, Vaughan TG, Egli A, Tschudin-Sutter S, Stadler T, Bonhoeffer S. 2022. The effect of sequencing and assembly on the inference of horizontal gene transfer on chromosomal and plasmid phylogenies. Philos Trans R Soc Lond B Biol Sci 377:20210245. https://doi.org/10.1098/rstb.2021.0245.

21. Campos ACC, Andrade NL, Ferdous M, Chlebowicz MA, Santos CC, Correal JCD, Lo Ten Foe JR, Rosa ACP, Damasco PV, Friedrich AW, Rossen JWA. 2018. Comprehensive molecular characterization of Escherichia coli isolates from urine samples of hospitalized patients in Rio de Janeiro, Brazil. Front Microbiol 9:243. https://doi.org/10.3389/fmicb.2018.00243.

22. Rogers BA, Ingram PR, Runnegar N, Pitman MC, Freeman JT, Athan E, Havers S, Sidjabat HE, Gunning E, De Almeida M, Styles K, Paterson DL. 2015. Sequence type 131 fimH30 and fimH41 subclones amongst Escherichia coli isolates in Australia and New Zealand. Int J Antimicrobial Agents 45:351–358. https://doi.org/10.1016/j.ijantimicag.2014.11.015.

23. Li B, Lu Y, Lan F, He Q, Li C, Cao Y. 2017. Prevalence and characteristics of ST131 clone among unselected clinical Escherichia coli in a Chinese university hospital. Antimicrob Resist Infect Control 6:118. https://doi.org/10.1186/s13756-017-0274-0.

24. Office of the High Commissioner for Diaspora Affairs. Armenian Diaspora Communities. http://diaspora.gov.am/en/diasporas. Retrieved 20 September 2022.

25. Soto SM. 2009. Human migration and infectious diseases. Clin Microbiol Infect 15:26–28. https://doi.org/10.1111/j.1469-0691.2008.02694.x.

26. Abdelrahim SS, Fouad M, Abdallah N, Ahmed RF, Zaki S. 2021. Comparative study of CTX-M-15 producing Escherichia coli ST131 clone isolated from urinary tract infections and acute diarrhoea. Infect Drug Resist 14: 4027–4038. https://doi.org/10.2147/IDR.S325669.

27. Zhong Y-M, Liu W-E, Liang X-H, Li Y-M, Jian Z-J, Hawkey PM. 2015. Emergence and spread of O16-ST131 and O25b-ST131 clones among faecal CTX-M-producing Escherichia coli in healthy individuals in Hunan Province, China. J Antimicrob Chemother 70:2223–2227. https://doi.org/10.1093/jac/dkv114.

28. Phan MD, Forde BM, Peters KM, Sarkar S, Hancock S, Stanton-Cook M, Ben Zakour NL, Upton M, Beatson SA, Schembri MA. 2015. Molecular characterization of a multidrug resistance IncF plasmid from the globally disseminated Escherichia coli ST131 clone. PLoS One 10:e0122369. https://doi.org/10.1371/journal.pone.0122369.

29. Whitmer GR, Moorthy G, Arshad M. 2019. The pandemic Escherichia coli sequence type 131 strain is acquired even in the absence of antibiotic exposure. PLoS Pathog 15:e1008162. https://doi.org/10.1371/journal.ppat.1008162.

30. Ludden C, Decano AG, Jamrozy D, Pickard D, Morris D, Parkhill J, Peacock SJ, Cormican M, Downing T. 2020. Genomic surveillance of Escherichia coli ST131 identifies local expansion and serial replacement of subclones. Microb Genom 6:e000352. https://doi.org/10.1099/mgen.0.000352.

31. Shaik S, Ranjan A, Tiwari SK, Hussain A, Nandanwar N, Kumar N, Jadhav S, Semmler T, Baddam R, Islam MA, Alam M, Wieler LH, Watanabe H, Ahmed N. 2017. Comparative genomic analysis of globally dominant ST131 clone with other epidemiologically successful extraintestinal pathogenic Escherichia coli (ExPEC) lineages. mBio 8:e01596-17. https://doi.org/10.1128/mBio.01596-17.

32. Shawa M, Furuta Y, Mulenga G, Mubanga M, Mulenga E, Zorigt T, Kaile C, Simbotwe M, Paudel A, Hang'ombe B, Higashi H. 2021. Novel chromosomal insertions of ISEcp1-blaCTX-M-15 and diverse antimicrobial resistance genes in Zambian clinical isolates of Enterobacter cloacae and Escherichia coli. Antimicrob Resist Infect Control 10:79. https://doi.org/10.1186/s13756-021-00941-8.

33. He S, Hickman AB, Varani AM, Siguier P, Chandler M, Dekker JP, Dyda F. 2015. Insertion sequence IS26 reorganizes plasmids in clinically isolated multidrug-resistant bacteria by replicative transposition. mBio 6:e00762-15. https://doi.org/10.1128/mBio.00762-15.

34. Rozwandowicz M, Brouwer MSM, Fischer J, Wagenaar JA, Gonzalez-Zorn B, Guerra B, Mevius DJ, Hordijk J. 2018. Plasmids carrying antimicrobial resistance genes in Enterobacteriaceae. J Antimicrob Chemother 73:1121–1137. https://doi.org/10.1093/jac/dkx488.

35. Hayashi M, Matsui M, Sekizuka T, Shima A, Segawa T, Kuroda M, Kawamura K, Suzuki S. 2020. Dissemination of IncF group F1:A2:B20 plasmid-harbouring multidrug-resistant Escherichia coli ST131 before the acquisition of blaCTX-M in Japan. J Glob Antimicrob Resist 23:456–465. https://doi.org/10.1016/j.jgar.2020.10.021.

36. Partridge SR, Zong Z, Iredell JR. 2011. Recombination in IS26 and Tn2 in the evolution of multiresistance regions carrying blaCTX-M-15 on conjugative IncF plasmids from Escherichia coli. Antimicrob Agents Chemother 55:4971–4978. https://doi.org/10.1128/AAC.00025-11.

37. Shropshire WC, Aitken SL, Pifer R, Kim J, Bhatti MM, Li X, Kalia A, Galloway-Peña J, Sahasrabhojane P, Arias CA, Greenberg DE, Hanson BM, Shelburne SA. 2021. IS26-mediated amplification of blaOXA-1 and blaCTX-M-15 with concurrent outer membrane porin disruption associated with de novo carbapenem resistance in a recurrent bacteraemia cohort. J Antimicrob Chemother 76:385–395. https://doi.org/10.1093/jac/dkaa447.

38. Harmer CJ, Moran RA, Hall RM. 2014. Movement of IS26-associated antibiotic resistance genes occurs via a translocatable unit that includes a single IS26 and preferentially inserts adjacent to another IS26. mBio 5:e01801-14. https://doi.org/10.1128/mBio.01801-14.

39. Harmer CJ, Lebreton F, Stam J, McGann PT, Hall RM. 2022. Mechanisms of IS26-mediated amplification of the aphA1 gene leading to tobramycin resistance in an Acinetobacter baumannii isolate. Microbiol Spectr 10:e02287-22. https://doi.org/10.1128/spectrum.02287-22.

40. Yang Q-E, Sun J, Li L, Deng H, Liu B-T, Fang L-X, Liao X-P, Liu Y-H. 2015. IncF plasmid diversity in multi-drug resistant Escherichia coli strains from animals in China. Front Microbiol 6:964. https://doi.org/10.3389/fmicb.2015.00964.

41. Massip C, Oswald E. 2020. Siderophore-microcins in Escherichia coli: determinants of digestive colonization, the first step toward virulence. Front Cell Infect Microbiol 10:381. https://doi.org/10.3389/fcimb.2020.00381.

42. Abraham S, Chapman TA, Zhang R, Chin J, Mabbett AN, Totsika M, Schembri MA. 2012. Molecular characterization of Escherichia coli strains that cause symptomatic and asymptomatic urinary tract infections. J Clin Microbiol 50:1027–1030. https://doi.org/10.1128/JCM.06671-11.

43. Massip C, Chagneau CV, Boury M, Oswald E. 2020. The synergistic triad between microcin, colibactin, and salmochelin gene clusters in uropathogenic Escherichia coli. Microbes Infect 22:144–147. https://doi.org/10.1016/j.micinf.2020.01.001.

44. Snyder JA, Haugen BJ, Buckles EL, Lockatell CV, Johnson DE, Donnenberg MS, Welch RA, Mobley HLT. 2004. Transcriptome of uropathogenic Escherichia coli during urinary tract infection. Infect Immun 72:6373–6381. https://doi.org/10.1128/IAI.72.11.6373-6381.2004.

45. Hancock V, Vejborg RM, Klemm P. 2010. Functional genomics of probiotic Escherichia coli Nissle 1917 and 83972, and UPEC strain CFT073: comparison of transcriptomes, growth and biofilm formation. Mol Genet Genomics 284:437–454. https://doi.org/10.1007/s00438-010-0578-8.

46. Alqasim A, Jaffal AA, Almutairi N, Alyousef AA. 2021. Comparative phenotypic characterization identifies few differences in the metabolic capacity between Escherichia coli ST131 subclones. Saudi J Biol Sci 28:762–769. https://doi.org/10.1016/j.sjbs.2020.11.008.

47. Alqasim A, Emes R, Clark G, Newcombe J, Ragione RL, McNally A. 2014. Phenotypic microarrays suggest Escherichia coli ST131 is not a metabolically distinct lineage of extra-intestinal pathogenic E. coli. PLoS One 9:e88374. https://doi.org/10.1371/journal.pone.0088374.

48. Hertz FB, Marvig RL, Frimodt-Møller N, Nielsen KL. 2022. In vitro relative fitness, in vivo intestinal colonization and genomic differences of Escherichia coli of ST131 carrying blaCTX-M-15. Front Microbiol 12:798473. https://doi.org/10.3389/fmicb.2021.798473.

49. Shin J, Ko KS. 2015. Effect of plasmids harbouring blaCTX-M on the virulence and fitness of Escherichia coli ST131 isolates. Int J Antimicrob Agents 46:214–218. https://doi.org/10.1016/j.ijantimicag.2015.04.012.

50. Cave R, Ter-Stepanyan MM, Kotsinyan N, Mkrtchyan HV. 2022. An emerging lineage of uropathogenic extended spectrum β-lactamase Escherichia coli ST127. Microbiol Spectr 10:e02511-22. https://doi.org/10.1128/spectrum.02511-22.

51. Mkrtchyan HV, Russell CA, Wang N, Cutler RR. 2013. Could public restrooms be an environment for bacterial resistomes? PLoS One 8:e54223. https://doi.org/10.1371/journal.pone.0054223.

52. European Committee on Antimicrobial Susceptibility Testing. 2021. Breakpoint tables for interpretation of MICs and zone diameters. version 11.0. https://www.eucast.org/ast_of_bacteria/previous_versions_of_documents.

53. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

54. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. 2020. Using SPAdes de novo assembler. Curr Protoc Bioinformatics 70:e102. https://doi.org/10.1002/cpbi.102.

55. Perešíni P, Boža V, Brejová B, Vinař T. 2021. Nanopore base calling on the edge. Bioinformatics 37:4661–4667. https://doi.org/10.1093/bioinformatics/btab528.

56. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595.

57. Johnson J, Soehnlen M, Blankenship HM. 2023. Long read genome assemblers struggle with small plasmids. Microb Genom 9:e001024. https://doi.org/10.1099/mgen.0.001024.

58. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res 27:722–736. https://doi.org/10.1101/gr.215087.116.

59. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. 5. Nat Biotechnol 37:540–546. https://doi.org/10.1038/s41587-019-0072-8.

60. Wick RR, Holt KE. 2021. Benchmarking of long-read assemblers for prokaryote whole genome sequencing. F1000Res 8:2138. https://doi.org/10.12688/f1000research.21782.4.

61. Vaser R, Šikić M. 2021. Time- and memory-efficient genome assembly with Raven. Nat Comput Sci 1:332–336. https://doi.org/10.1038/s43588-021-00073-4.

62. Chen Y, Nie F, Xie S-Q, Zheng Y-F, Dai Q, Bray T, Wang Y-X, Xing J-F, Huang Z-J, Wang D-P, He L-J, Luo F, Wang J-X, Liu Y-Z, Xiao C-L. 2021. Efficient assembly of nanopore reads via highly accurate and intact error correction. Nat Commun 12:60. https://doi.org/10.1038/s41467-020-20236-7.

63. Hu J, Fan J, Sun Z, Liu S. 2020. NextPolish: a fast and efficient genome polishing tool for long-read assembly. Bioinformatics 36:2253–2255. https://doi.org/10.1093/bioinformatics/btz891.

64. Wick RR, Judd LM, Cerdeira LT, Hawkey J, Méric G, Vezina B, Wyres KL, Holt KE. 2021. Trycycler: consensus long-read assemblies for bacterial genomes. Genome Biol 22:266. https://doi.org/10.1186/s13059-021-02483-z.

65. Wick RR, Holt KE. 2022. Polypolish: short-read polishing of long-read bacterial genome assemblies. PLoS Comput Biol 18:e1009802. https://doi.org/10.1371/journal.pcbi.1009802.

66. Zimin AV, Salzberg SL. 2020. The genome polishing tool POLCA makes fast and accurate corrections in genome assemblies. PLoS Comput Biol 16:e1007981. https://doi.org/10.1371/journal.pcbi.1007981.

67. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, Parkhill J, Harris SR. 2015. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. Nucleic Acids Res 43:e15. https://doi.org/10.1093/nar/gku1196.

68. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32:268–274. https://doi.org/10.1093/molbev/msu300.

69. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. 2018. UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol 35:518–522. https://doi.org/10.1093/molbev/msx281.

70. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods 14:587–589. https://doi.org/10.1038/nmeth.4285.

71. Letunic I, Bork P. 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res 49:W293–W296. https://doi.org/10.1093/nar/gkab301.

72. Rambaut A, Lam TT, Max Carvalho L, Pybus OG. 2016. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). Virus Evol 2:vew007. https://doi.org/10.1093/ve/vew007.

73. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.

74. Bessonov K, Laing C, Robertson J, Yong I, Ziebell K, Gannon VPJ, Nichani A, Arya G, Nash JHE, Christianson S. 2021. ECTyper: in silico Escherichia coli serotype and species prediction from raw and assembled whole-genome sequence data. Microb Genom 7:000728. https://doi.org/10.1099/mgen.0.000728.

75. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, Huynh W, Nguyen A-LV, Cheng AA, Liu S, Min SY, Miroshnichenko A, Tran H-K, Werfalli RE, Nasir JA, Oloni M, Speicher DJ, Florescu A, Singh B, Faltyn M, Hernandez-Koutoucheva A, Sharma AN, Bordeleau E, Pawlowski AC, Zubyk HL, Dooley D, Griffiths E, Maguire F, Winsor GL, Beiko RG, Brinkman FSL, Hsiao WWL, Domselaar GV, McArthur AG. 2020. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res 48:D517–D525.

76. Liu Y-F, Yan J-J, Lei H-Y, Teng C-H, Wang M-C, Tseng C-C, Wu J-J. 2012. Loss of outer membrane protein C in Escherichia coli contributes to both antibiotic resistance and escaping antibody-dependent bactericidal activity. Infect Immun 80:1815–1822. https://doi.org/10.1128/IAI.06395-11.

77. Choi U, Lee C-R. 2019. Distinct roles of outer membrane porins in antibiotic resistance and membrane integrity in Escherichia coli. Front Microbiol 10:953. https://doi.org/10.3389/fmicb.2019.00953.

78. Roer L, Tchesnokova V, Allesøe R, Muradova M, Chattopadhyay S, Ahrenfeldt J, Thomsen MCF, Lund O, Hansen F, Hammerum AM, Sokurenko E, Hasman H. 2017. Development of a web tool for Escherichia coli subtyping based on fimH alleles. J Clin Microbiol 55:2538–2543. https://doi.org/10.1128/JCM .00737-17.

79. Jolley KA, Bray JE, Maiden MCJ. 2018. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. Wellcome Open Res 3:124. https://doi.org/10.12688/wellcomeopenres.14826.1.

80. Villa L, García-Fernández A, Fortini D, Carattoli A. 2010. Replicon sequence typing of IncF plasmids carrying virulence and resistance determinants. J Antimicrob Chemother 65:2518–2529. https://doi.org/10.1093/jac/dkq347.

81. Robertson J, Nash JHE. 2018. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. Microb Genom 4: e000206. https://doi.org/10.1099/mgen.0.000206.

82. Carattoli A, Hasman H. 2020. PlasmidFinder and In Silico pMLST: identification and typing of plasmid replicons in whole-genome sequencing (WGS). Methods Mol Biol 2075:285–294. https://doi.org/10.1007/978-1-4939-9877-7_20.

83. Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M. 2006. ISfinder: the reference centre for bacterial insertion sequences. Nucleic Acids Res 34:D32–D36. https://doi.org/10.1093/nar/gkj014.

84. Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, Villa L, Møller 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. https://doi.org/10.1128/AAC.02412-14.

85. Liu M, Li X, Xie Y, Bi D, Sun J, Li J, Tai C, Deng Z, Ou H-Y. 2019. ICEberg 2.0: an updated database of bacterial integrative and conjugative elements. Nucleic Acids Res 47:D660–D665. https://doi.org/10.1093/nar/gky1123.

86. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 31:3691–3693. https://doi.org/10.1093/bioinformatics/btv421.

87. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34:3094–3100. https://doi.org/10.1093/bioinformatics/bty191.

88. Ondov BD, Starrett GJ, Sappington A, Kostic A, Koren S, Buck CB, Phillippy AM. 2019. Mash Screen: high-throughput sequence containment estimation for genome discovery. Genome Biol 20:232. https://doi.org/10.1186/s13059-019-1841-x.

89. Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK. 2016. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. Anal Methods 8:12–24. https://doi.org/10.1039/C5AY02550H.