

# Genomic profiling of Escherichia coli isolates from bacteraemia patients: a 3-year cohort study of isolates collected at a Sydney teaching hospital

**Author:**

Hastak, Priyanka; Cummins, Max L; Gottlieb, Thomas; Cheong, Elaine; Merlino, John; Myers, Garry SA; Djordjevic, Steven P; Chowdhury, Piklu Roy

**Publication details:**

Microbial Genomics

v. 6

Chapter No. 5

Medium: Print-Electronic

pp. e000371 - e000371

2057-5858 (ISSN)

**Publication Date:**

2020-05-01

**Publisher DOI:**

<https://doi.org/10.1099/mgen.0.000371>

Downloaded from [http://hdl.handle.net/1959.4/unsworks\\_85505](http://hdl.handle.net/1959.4/unsworks_85505) in <https://unsworks.unsw.edu.au> on 2024-05-18

# Genomic profiling of *Escherichia coli* isolates from bacteraemia patients: a 3-year cohort study of isolates collected at a Sydney teaching hospital

Priyanka Hastak<sup>1,2</sup>, Max L. Cummins<sup>1</sup>, Thomas Gottlieb<sup>3,4</sup>, Elaine Cheong<sup>3</sup>, John Merlino<sup>3,4</sup>, Garry S. A. Myers<sup>1</sup>, Steven P. Djordjevic<sup>1,2</sup> and Piklu Roy Chowdhury<sup>1,2,\*</sup>

## Abstract

This study sought to assess the genetic variability of *Escherichia coli* isolated from bloodstream infections (BSIs) presenting at Concord Hospital, Sydney during 2013–2016. Whole-genome sequencing was used to characterize 81 *E. coli* isolates sourced from community-onset (CO) and hospital-onset (HO) BSIs. The cohort comprised 64 CO and 17 HO isolates, including 35 multidrug-resistant (MDR) isolates exhibiting phenotypic resistance to three or more antibiotic classes. Phylogenetic analysis identified two major ancestral clades. One was genetically diverse with 25 isolates distributed in 16 different sequence types (STs) representing phylogroups A, B1, B2, C and F, while the other comprised phylogroup B2 isolates in subclades representing the ST131, ST73 and ST95 lineages. Forty-seven isolates contained a class 1 integron, of which 14 carried *bla*<sub>CTX-M-3</sub> gene. Isolates with a class 1 integron carried more antibiotic resistance genes than isolates without an integron and, in most instances, resistance genes were localized within complex resistance loci (CRL). Resistance to fluoroquinolones could be attributed to point mutations in chromosomal *parC* and *gyrB* genes and, in addition, two isolates carried a plasmid-associated *qnrB4* gene. Co-resistance to fluoroquinolone and broad-spectrum beta-lactam antibiotics was associated with ST131 (HO and CO), ST38 (HO), ST393 (CO), ST2003 (CO) and ST8196 (CO and HO), a novel ST identified in this study. Notably, 10/81 (12.3%) isolates with ST95 (5 isolates), ST131 (2 isolates), ST88 (2 isolates) and a ST540 likely carry IncFII–IncFIB plasmid replicons with a full spectrum of virulence genes consistent with the carriage of ColV-like plasmids. Our data indicate that IncF plasmids play an important role in shaping virulence and resistance gene carriage in BSI *E. coli* in Australia.

## DATA SUMMARY

The genome sequences of the isolates included in this paper have been made accessible via GenBank, under project ID PRJNA480723. A list of individual accession numbers for the assembled genomes is provided in File S1. The genome sequence of the novel *E. coli* ST8196 was registered in

Enterobase and is searchable using the sequence type numbers and Uber strain IDs ESC\_IA6781AA and ESC\_IA7243AA. All online databases were accessed through portals on the Centre Of Genomic Epidemiology website (<http://www.genomicepidemiology.org/>).

Received 14 September 2019; Accepted 03 April 2020; Published 06 May 2020

**Author affiliations:** <sup>1</sup>The ithree institute, University of Technology Sydney, City Campus, Ultimo, NSW 2007, Australia; <sup>2</sup>Australian Centre for Genomic Epidemiological Microbiology, University of Technology Sydney, PO Box 123, Broadway, NSW 2007, Australia; <sup>3</sup>Department of Microbiology and Infectious Diseases, Concord Hospital and NSW Health Pathology, Hospital Road, Concord 2139, NSW, Australia; <sup>4</sup>Faculty of Medicine, University of Sydney, NSW Australia.

**\*Correspondence:** Piklu Roy Chowdhury, piklu.bhattacharya@uts.edu.au

**Keywords:** *Escherichia coli*; bacteraemia; blood stream infections; ESBL; fluoroquinolone resistance; multidrug resistance; class 1 integrons; IncF plasmids; IS26; *bla*<sub>CTX-M</sub>; *E. coli*ST8196; *qnrB4*.

**Abbreviations:** AGAR, Australian Group for Antimicrobial Resistance; AMR, antimicrobial resistance; BSI, blood stream infections; CO, community onset; CRL, complex resistance loci; EIEC, Enteroinvasive *Escherichia coli*; ESBL, extended spectrum beta-lactamase; ETEC, Enterotoxigenic *Escherichia coli*; ExPEC, Extra-intestinal pathogenic *Escherichia coli*; GNSOP, gram-negative bacteraemia surveillance programme; HO, hospital onset; MDR, multidrug resistance/resistant; ST, sequence type; UPEC, Uropathogenic *Escherichia coli*; UTI, urinary tract infections; VAG, virulence associated genes; WGS, whole genome sequencing.  
GenBank Project ID PRJNA480723.

**Data statement:** All supporting data, code and protocols have been provided within the article or through supplementary data files. Six supplementary tables and one supplementary figure is available with the online version of this article.

000371 © 2020 The Authors



This is an open-access article distributed under the terms of the Creative Commons Attribution NonCommercial License.

## INTRODUCTION

*Escherichia coli* is a major cause of bloodstream infections (BSIs), dominating the list of Gram-negative pathogens associated with bacteraemia and sepsis globally [1]. In 2017, treatment costs for patients with sepsis in Australian intensive care units were estimated to be at least USD \$846 million [2]. In that year, 55.2% of Gram-negative BSIs were caused by *E. coli*, of which 83.6% were from community and 16.4% from hospitalized patients [3]. Nationwide, the incidence of multidrug resistance associated with *E. coli* BSIs increased from 11.8% in 2013 [4] to 21.9% in 2017 [3]. This was also accompanied by an increase in the frequency of the isolation of multidrug-resistant (MDR) *E. coli* displaying resistance to fluoroquinolones and extended-spectrum beta-lactams [3].

Extra-intestinal pathogenic *E. coli* (ExPEC) are usually associated with BSIs and sepsis. ExPEC causing human infections are often characterized by the presence of P and S fimbriae that facilitate human cell attachment, toxins and cytotoxic factors that enable invasion and subsequent damage, iron acquisition systems, serum resistance proteins [5, 6], capsular polysaccharides and lipopolysaccharides (LPS) that support survival in the host. *E. coli* BSIs predominantly have a community onset, with urinary tract infections (UTIs) being the most frequently reported principal clinical manifestation [3, 7]. Uropathogenic *E. coli* (UPEC) are a subgroup of ExPEC that are adept at colonizing the urethra and ascending into the urinary bladder causing cystitis. A 3-month sentinel surveillance of *E. coli* bacteraemia cases in the UK in 2012/2013, which included 35 hospitals, linked 51.2% of the cases to UTI [8]. The first-line drugs used to treat UTI in Australia include trimethoprim, amoxicillin/clavulanate and cephalexin. Genes conferring resistance to trimethoprim are often present as gene cassettes on class 1 integrons, a globally disseminated mobile genetic element best known for its role in the rapid evolution of MDR pathogens [9–11]. Within enterobacterial genomes, class 1 integrons are frequently detected in regions where resistance genes cluster together, forming resistance loci. Hence they serve as a reliable proxy for the identification of multiple drug resistant genomes within cohorts of bacteria collected from clinical settings [12], food animals [13–15] and the environment [16]. Increasing resistance to aminoglycosides, extended-spectrum beta-lactams, fluoroquinolones and trimethoprim is a major threat to the treatment of *E. coli* infections, including BSIs [17–21]. Plasmids play a major role in driving the evolution of pathogens resistant to these antibiotics through the rapid dispersal of virulence and antibiotic resistance genes [22, 23] within enterobacteriales. Examples of plasmid-mediated rapid pathogen evolution include the establishment of enterotoxigenic *E. coli* (ETEC) [24, 25] and enteroinvasive *E. coli* (EIEC) [26] pathotypes, which cause watery diarrhoea and bacillary (often bloody) dysentery, respectively.

Globally, the most predominant ExPEC sequence types (STs) associated with UTI include ST69, ST73, ST95 and ST131 [27, 28]. Specifically, the ST131 lineage is recognized as a significant aetiological agent of UTI, pyelonephritis and

### Impact Statement

We report on the genomic surveillance of 81 *Escherichia coli* isolates collected from patients treated for bloodstream infections (BSIs) at the Concord Repatriation Hospital in Sydney, Australia between January 2013 and March 2016. Our data describe the population dynamics and highlight a segregation between virulent and drug-resistant clonal lineages in the cohort, akin to trends reported recently in a study on a cohort of more than 1500 invasive *E. coli* isolates collected over 11 years in the UK. Our study identifies a novel sequence type of *E. coli*, ST8196, within the ST131 clonal cluster, and the possible role of large ColV-like IncF plasmids in shaping the virulence and resistance profiles of isolates in this cohort. We additionally present compelling evidence on the active role played by IS26 in the ongoing evolution and clustering of antimicrobial resistance genes in *E. coli* from both community-onset (CO) and hospital-onset (HO) BSIs. Collectively, our results highlight why surveillance protocols that target the mobilome are needed to better understand features that distinguish CO and HO infections.

urosepsis [29, 30]. The wide distribution of ST131, together with relatively higher numbers of drug resistance and virulence genes, has made it the most well-studied ExPEC lineage [31, 32]. The emergence of MDR BSIs is often associated with isolates that have acquired genes encoding extended-spectrum beta-lactamases (ESBLs), particularly *bla*<sub>CTX-M</sub>. There are over 100 allelic variants of *bla*<sub>CTX-M</sub> [33–35] and different alleles predominate in *E. coli*, depending on several factors, including geographical location [36]. BSIs in Argentina, South Africa, Turkey, Belgium, China, Taiwan and Australia are typically caused by *E. coli* that carry *bla*<sub>CTX-M-2</sub>, *bla*<sub>CTX-M-3</sub> and *bla*<sub>CTX-M-15</sub> [37–39].

Surveillance of antibiotic-resistant pathogens has provided important insights into the distribution and prevalence of antibiotic resistance in Australia [40], and has highlighted rising rates of ESBL-producing and fluoroquinolone-resistant MDR *E. coli* causing BSIs [3, 7, 41]. Whole-genome sequencing (WGS) of pathogens causing BSIs facilitates the identification of AMR gene carriage and virulence-associated genes (VAGs) and links their carriage with ST, phylogroup and e-serotype. It also provides insights into the genetic context of AMR genes and VAGs and sheds light on the associations of AMR genes with mobile genetic elements. In addition to these attributes, WGS enables effective microbial source tracking.

In this report, we present an analysis of *E. coli* whole-genome sequences collected from patients with bacteraemia (BSI) attending Concord Hospital, Sydney between January 2013 and March 2016. We investigate the genetic diversity of *E. coli* that cause BSIs, describe the genetic context of the resistance

gene carried by these isolates and identify genetic elements that disseminate AMR genes and VAGs.

## METHODS

### *E. coli* collection and growth conditions

Between January 2013 and March 2016, 406 *E. coli* isolates were collected by the Microbiology Department at Concord Repatriation Hospital in Sydney and stored as part of the Australian Group for Antimicrobial Resistance (AGAR) Gram-negative bacteraemia surveillance programme (GNSOP). We selected 81 isolates representative of the spectrum of phenotypic antibiotic resistance profiles for whole-genome sequence analysis. Clinical metadata, including collection year, likely source of infection and phenotypic resistance profile, are presented in File S1 (available in the online version of this article). A 48 h post-admission cut-off for sample collection was used to distinguish community-onset (CO) and hospital-onset (HO) infections. Isolates were sub-cultured and grown at 37 °C for 18 h on lysogeny broth (LB) and the pellets were used for purification of DNA.

### Antimicrobial resistance phenotyping

Antibiotic susceptibility profiles were generated using the Advanced Expert System (AES) software designed for VITEK 2 (BioMerieux). Breakpoints were interpreted using European Committee on Antimicrobial Susceptibility Testing (EUCAST) guidelines [42] at the Microbiology Laboratory, Concord Hospital. The panel of antibiotics tested comprised ampicillin (AMP), amoxicillin (AMC), ciprofloxacin (CIP), cefazolin (CFZ), cefotaxime (CTX), ceftriaxone (CRO), ceftazidime (CAZ), ceftiofloxacin (FOX), trimethoprim (TRI), gentamicin (GEN), tobramycin (TOB), norfloxacin (NOR), ticarcillin/clavulanic acid (TCC), piperacillin/tazobactam (TZP), trimethoprim/sulfamethoxazole (SXT) and amikacin (AMK). Synergy tests were used to phenotypically identify isolates producing ESBL enzymes (File S1). For the purposes of this study, multi-drug resistance was defined as resistance to  $\geq$  three classes of antibiotics [43].

### Genomic DNA extraction and sequencing library preparation

Genomic DNA was extracted from 2 ml of overnight cultures using the ISOLATE II Genomic DNA kit (Biolone, Australia), following the manufacturer's protocol, and quantified using the Qubit fluorimeter and dsDNA HS Assay kit (Thermo Fisher Scientific, Australia). WGS libraries were prepared from 2 ng of gDNA using the Illumina Nextera XT DNA kit following previously published protocols [44].

### WGS and assembly

WGS was performed using the Illumina HiSeq 2500 v4 and the rapid PE150 mode (Illumina, San Diego, CA, USA). Based on FastQC reports, mean read lengths between 120 nt and 125 nt (with Phred scores  $\geq$ 30) were used for genome assembly using the A5-miSeq pipeline (version 20150522) [45]. The number of scaffolds per genome varied from 30

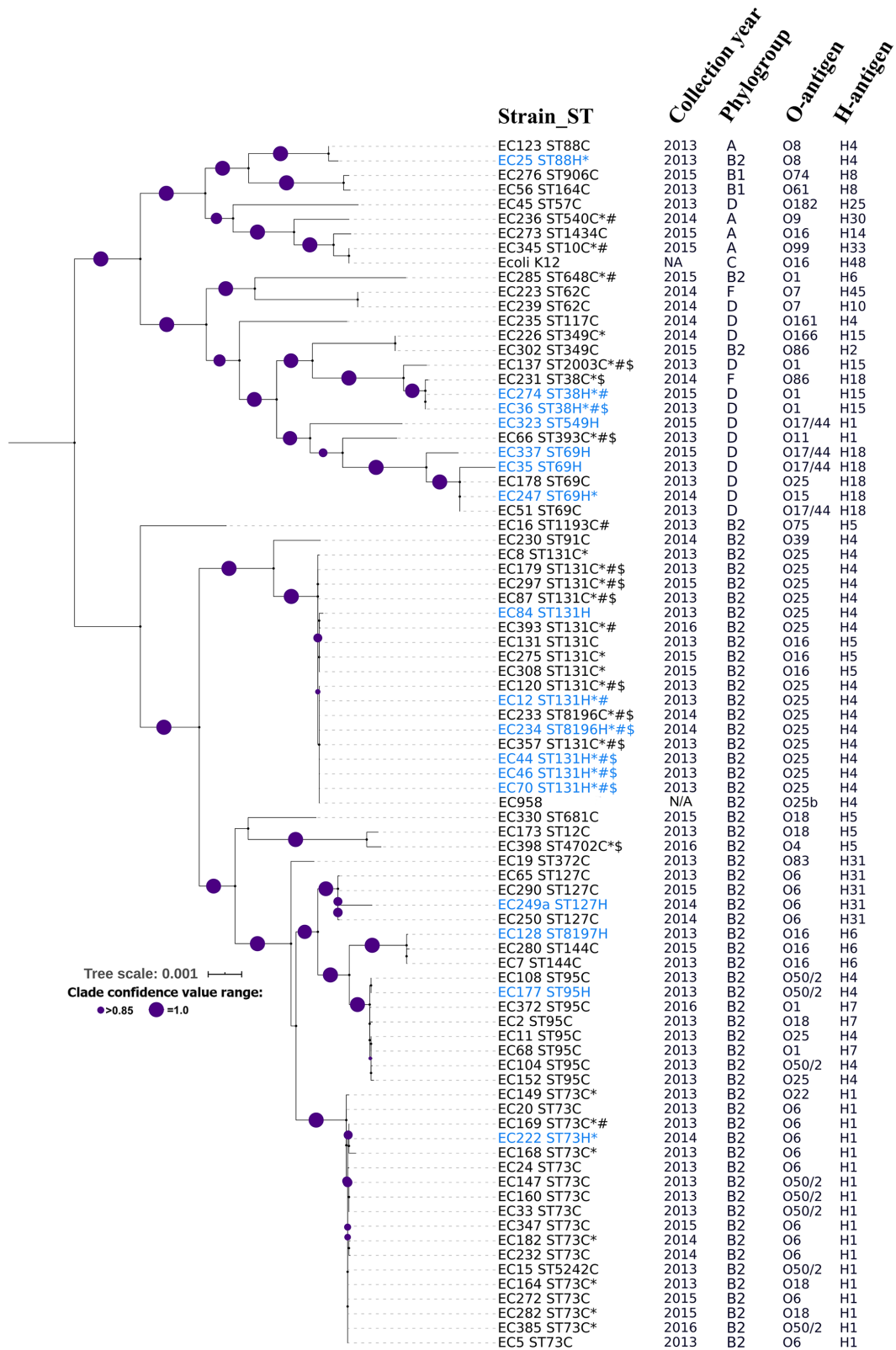
to 200, with median sequence coverage ranging between 23 $\times$  and 54 $\times$  (File S2). Preliminary genome annotations were performed on RASTtk (<http://rast.nmpdr.org/>) [46] and regions of interest were manually curated using online versions of BLASTn and iterative BLASTp searches. The assembled genome sequences are available in GenBank under project ID PRJNA480723. The BioSample accession numbers are listed in File S2.

### Gene identification and multi-locus sequence type (MLST) analysis

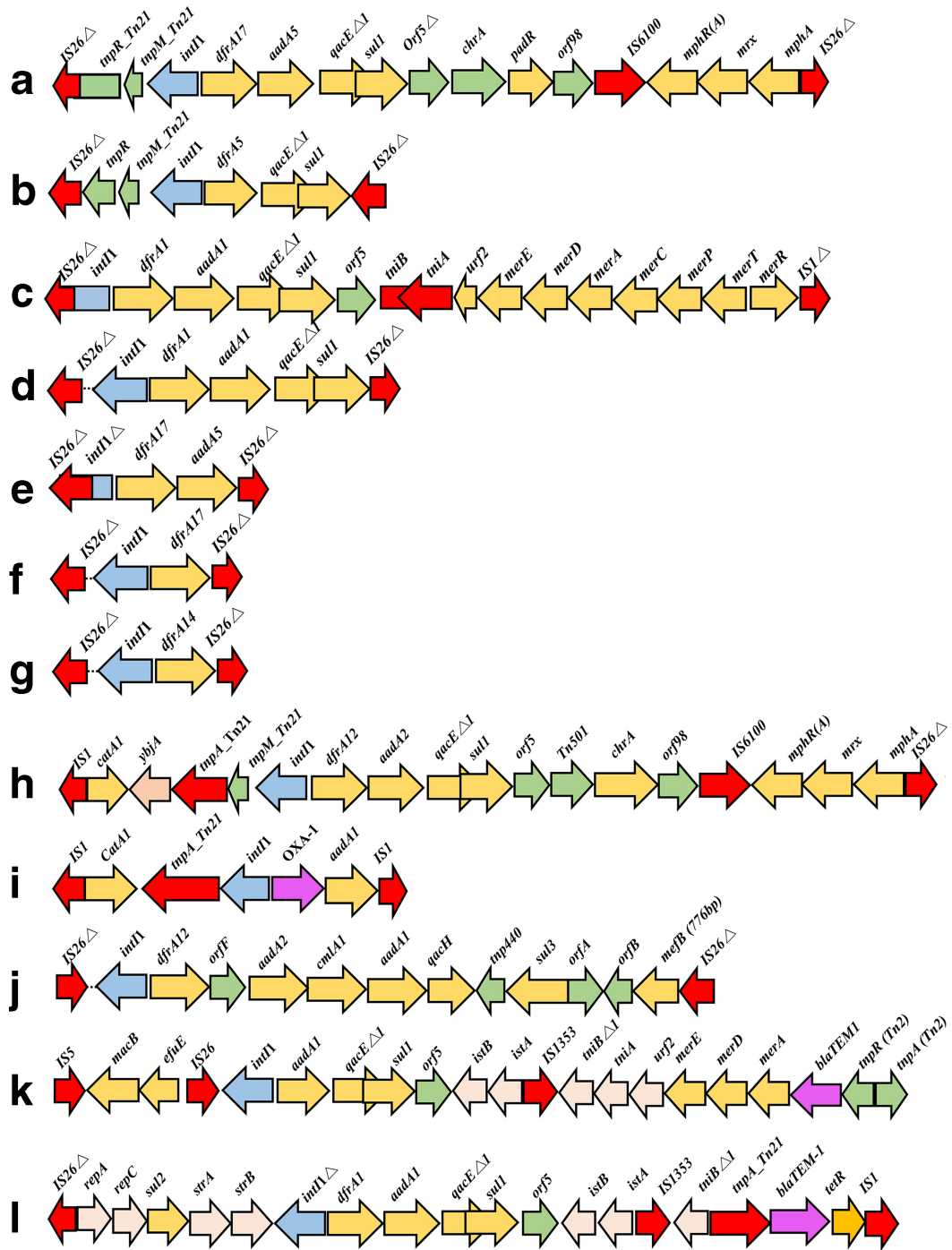
The Antimicrobial Resistance Identification By Assembly (ARIBA) pipeline (v2.6.1) [47] was used to rapidly detect the presence of the class 1 integrase gene *intI1*, antimicrobial resistance genes (ARGs), virulence-associated genes (VAGs), specific insertion elements, O- and H-antigen genes and Clermont phylogroups. The output was filtered strictly for hits that had 100% query coverage and  $\geq$ 99% sequence identity. Genomes that did not return a match based on our stringent parameters were individually assessed either by BLASTn analysis (*intI1* gene) or via portals on the Centre of Genomic Epidemiology website using assembled sequences (<http://www.genomicsepidemiology.org/>) of respective genomes. Raw Illumina reads were mapped to pCERC-4 (accession no. KU578032) and pSDJ2009-52F (accession no. MH195200.1) to assess presence of genetic markers of the IncFII–IncFIB–ColV hybrid plasmids using published protocols [13]. The read mapping exercise was complemented with BLAST analysis of pCERC-4 against assembled genome sequences to present summary of findings presented in the manuscript. Genomes exhibiting  $\geq$ 95% sequence identity over 100% query sequence of (i) a full complement of the colicin-V operon, (ii) virulence genes on pCERC-4 and (iii) linkage (defined by presence on the same genomic scaffold) of at least one of the *repA* genes on pCERC-4 with two or more Col-V-specific virulence gene/s, exhibiting specified identity cutoffs were identified as candidate genomes that likely host a Col-V plasmid. MLST analysis was performed using the Achtman *E. coli* MLST scheme on the PubMLST database (<http://pubmlst.org/>) and the plasmid replicons were identified using the Plasmid Finder database (<https://cge.cbs.dtu.dk/services/PlasmidFinder/>) [48].

### Phylogenetic analysis

Reference gene-based maximum-likelihood phylogeny analysis (using 37 prokaryotic marker genes) was performed using PhyloSift (v1.0.1) [49]. FastTree version 2.1.8.c was used to construct a phylogenetic tree (<http://www.microbesonline.org/fasttree/FastTree-2.1.8.c>) and visualized using FigTree v1.4.2 [44]. The reference genomes used in the initial analysis were *E. coli* K12 (U00096.2) and *E. coli* EC958 (NZ\_HG941718.1) and *Klebsiella pneumoniae* subsp. *pneumoniae* MGH78578 (CP000647.1) was used as an outlier. To increase the resolution of the branches, the outer layer was removed from the final tree presented in this paper (Figs 1 and 2). The tree figures were edited using iTol



**Fig. 1.** A mid-point rooted maximum-likelihood phylogenetic tree of isolates included in the cohort. Panels on the right-hand side indicate their year of isolation, *in silico* phylogrouping, O-antigen and H-antigen profiles. The tree includes two reference genomes, *E. coli* K12 and EC958. The year of isolation was not available for the two reference genomes and hence is labelled as NA NA in the metadata column. The strain labels coloured in blue indicate isolates from hospital-acquired infections. MDR strains are marked with an asterisk (\*), ESBL producers are marked with a dollar sign (\$) and fluoroquinolone-resistant isolates are marked with a hash (#).



**Fig. 2.** Diagrammatic representation of class 1 integron-associated complex resistance loci identified in this study. The terminal IS26s (red arrows) are mostly fragmented, and their relative orientation depicted in the diagram is based on analysis of the partial IS26 sequence. Figure not to scale.

online tree management software (<https://itol.embl.de/>) and virulence heatmaps were created using the Burrows–Wheeler Aligner (BWA) 0.7.17, samtools 0.1.18 and ggplot2 in the R package [50]. The scripts and commands utilized for the heatmap are available on Github (<https://github.com/maxlcummins/APEC-MGEN-2018>).

## RESULTS

### Isolate metadata

Clinical metadata corresponding to the 81 deidentified bacterial samples (File. S1) indicated that 37 isolates (45.5%) were associated with UTIs and most of these (64/81; 79%) were

recovered from CO infections. The cohort included 35 MDR isolates of which 13 exhibited co-resistance to fluoroquinolones and ESBLs. Assembly statistics (File S2) indicated  $\geq 23\times$  coverage.

### Phylogeny and ST distribution

Thirty different *E. coli* STs (File. S1) were identified within this BSI cohort. Phylogenetic analysis revealed two major ancestral clusters (Fig. 1). The upper major clade contained 25 isolates distributed in 16 STs (Fig. 1). Fifty-six isolates that clustered in the lower major clade comprised 15 STs, including globally recognized *E. coli* clonal groups ST73 (17 isolates), ST131 (15 isolates), ST95 (8 isolates) and ST127 (4 isolates). Isolates in the upper major clade of the phylogenetic tree contained representative of Clermont phylogroups B2, D, F, A and B1, while all isolates in the lower major clade of the phylogenetic tree, irrespective of ST type and O and H surface antigens typed as phylogroup B2, known to be associated with pathogenic strains. The relative abundance of CO infections within clinically significant STs were ST73 (16 of 17), ST131 (10 of 15), ST95 (7 of 8) and ST69 (2 of 5) was noteworthy. The resistance genes, class 1 integron-associated genes and plasmid replicon types in the entire cohort are presented in Table 1.

*E. coli* STs ST8196 (two isolates) and ST8197 (one isolate) are novel and were registered in Enterobase (<https://enterobase.warwick.ac.uk/species/index/ecoli>). ST8196 is closely related to the ST131-O25:H4 sub-clade (Fig. 1) and is a new member of the ST131 clonal complex. The ST8196 isolates, EC233 and EC234, were MDR and were isolated from blood specimens of separate patients within a 24 h period. EC233 was from a patient with CO UTI, while EC234 was from a patient with a hospital-onset infection. Both isolates displayed phenotypic resistance to fluoroquinolones and produced ESBLs. The ST8197 isolate EC128 is closely related to members of the ST144 subclade (Fig. 1), a ST that shares the closest common ancestor with ST95. Isolate EC128\_ST8197 was resistant to ampicillin and trimethoprim/sulfamethoxazole and susceptible to all other antibiotics tested.

Most (21; 60%) of the 35 MDR isolates clustered within the ST131 (13 isolates) and ST73 (8 isolates) subclades (Fig. 1). Genes responsible for ESBL production and fluoroquinolone resistance phenotypes in the MDR -cohort were not restricted to isolates in the ST131 sub-clade, but also present in ST38 (HO isolate EC36), ST393 (CO isolate EC66), ST2003 (CO isolate EC137) and ST8196 (EC233 from CO and EC234 from HO) representatives.

### Genetic context of class 1 integrons and resistance genes

Of the 81 isolates, 47 (58%) carried a class 1 integron (Table 1). Aminoglycoside adenylyl transferase genes *aadA1*, *aadA2* (streptomycin resistance) and *aadA5* (spectinomycin resistance), dihydrofolate reductase gene alleles *dfrA1*, *dfrA7*, *dfrA14*, *dfrA14* and *dfrA17*, encoding resistance to trimethoprim, *cmlA1* (encoding chloramphenicol

resistance) and *bla*<sub>OXA-1</sub> (encoding host-specific low-level resistance to ESBL antibiotics), were most frequently present in cassette arrays associated with class 1 integrons in this collection (Table 1). The IncX4 replicon was identified in five isolates. Sixteen isolates harboured the *aadA5-dfrA17* [Fig. 2a (14), e(2)] cassette array, while others comprised different combinations of the gene cassettes listed above and in Table 1. Genomic scaffolds containing class 1 integrons often carried other clustered resistance genes (Fig. 2) interspersed with different insertion elements. Integrons bearing the *aadA5-dfrA17* cassette array were almost always linked to a macrolide resistance module encoded by *mphA* and *mphR* genes beyond the 3'-CS (conserved segment) of the class 1 integron, forming a complex resistance locus (CRL) (Fig. 2a). IS26 sequences flanked the majority of scaffolds carrying a class 1 integron (Fig. 2). Most isolates (64/81; 79%), irrespective of the resistance gene carriage, carried multiple copies of IS26 (File S1), which likely contributed to fragmentations in genome assemblies.

The genes encoding resistance to ESBL were restricted to three allelic variants of *bla*<sub>CTX-M</sub>: *bla*<sub>CTX-M-14</sub>, *bla*<sub>CTX-M-15</sub> and *bla*<sub>CTX-M-27</sub>. In most instances, *bla*<sub>CTX-M-15</sub> genes were associated with *ISEcp1* (File S1). In EC137\_ST2003 and EC398\_ST4702, *bla*<sub>CTX-M-14</sub> was associated with *ISEcp1b*. Isolate EC345\_ST10 had a class 1 integron with a *bla*<sub>OXA-1</sub> gene cassette (Fig. 2i), but with no detectable enzymatic activity. Fluoroquinolone resistance in most isolates was attributed to mutations in *gyrB* and *parC* chromosomal genes except in isolates EC46\_ST131, EC44\_ST131 and EC137\_ST2003, which additionally had acquired *aac(6')-Ib-cr*, a gene known to confer low-level resistance to fluoroquinolones and aminoglycosides. Fluoroquinolone resistance gene *qnrB4* was identified in isolates EC233\_ST8196 and EC234\_ST8196.

### *E. coli* phylogroups and virulence gene profiling

Eighty-one genomes were screened against a panel of 65 VAGs commonly found in pathogenic *E. coli* (Fig. 3, Files S3 and S4). The relative abundance of VAGs was highest in ST95 and ST73 isolates, followed by ST131 (Fig. 3, File. S4). In congruence with recent reports on co-relation between phylogroups and virulence gene profiles [51, 52], isolates belonging to Clermont phylogroups A, B1 and F carried fewer VAGs than the isolates in phylogroups B2 and D (File S4). The *fimH*, *yeiT*, *fyuA*, *irp2*, *sitA* and *iss* genes were identified in over 80% of the isolates in our cohort (File S3). Only four isolates, EC104, EC2, EC230 and EC19 (all phylogroup B2), carried *ibeA*. Of the 2 iron acquisition genes, *iucD* and *irp2*, *iucD* was present in 52 isolates (8 phylogroup D and 1 each of phylogroups A and D, with the rest B2), and 72 tested positive for *irp2* (all 68 phylogroup B2 and 2 each of A and F). A previous single hospital study of *E. coli* from BSIs also highlighted an abundance of *fimH*, *papC*, *iha* and *ibeA* genes [53].

The distribution of the six ExPEC-specific protectins, namely *kpsMT II*, *kpsMT III*, *traT*, *ompT*, *iss* and *cavC*, and two toxins, *hly* and *cnf1*, was also determined. Forty isolates tested

**Table 1.** Resistance genes and plasmid profile of isolates sequenced and analysed in this study.

Isolate descriptor	Resistance genes in the genome	Class 1 integron-associated genes	ESBL genes	Fluoroquinolone resistance genes	Plasmid replicons
EC70_ST131H * \$ #	<i>aac-3-lla, aacA4, aadA5, dfrA17, sul1, TEM-1B</i>	<i>aadA5, dfrA17, mph(A), sul1</i>	CTX-M-15, OXA-1	<i>gyrA(S83L, D87N); parC(S80I, E84V)</i>	IncFII
EC137_ST2003C * \$ #	<i>aac-3-Ib, aac(6')-1b-cr, aacA4, cmlA1, dfrA1, mph(A), sul1, tet(B), TEM-1B</i>	<i>aacA4, cmlA1</i>	CTX-M-14	<i>gyrA(S83L, D87N); parC(S80I), aac(6')-1b-cr</i>	IncFII, IncFIB, IncFIA, Col156, pColBS512
EC234_ST8196H * \$ #	<i>qnrB4, aac-3-lla, aadA5, dfrA17, mph(A), sul1, blaDHA-1</i>	<i>aadA5, dfrA17, sul1, mph(A)</i>	CTX-M-15	<i>gyrA(S83L, D87N); parC(S80I, E84V), qnrB4</i>	IncFII, IncI1, Col156, pColBS512
EC345_ST10C * #	<i>aadA1, catA1, tet(B)</i>	<i>OXA1, aadA1</i>	OXA-1	<i>gyrA(S83L, D87N); parC(S80I, E84V)</i>	IncFII
EC385_ST73C *	<i>aph-3-Ia, TEM-1B, dfrA17, mph(A), strA/B, sul2, tet(B)</i>	<i>dfrA17</i>	–	<i>gyrA(S83L)</i>	IncFII, IncFIB, Col156, pColBS512
EC308_ST131C *	<i>catA1, dfrA7, sul1, TEM-1B</i>	<i>catA1, dfrA7</i>	–	–	IncFII, IncFIA, IncFIB, Col156
EC236_ST540C * #	<i>aph-3-la, dfrA5, sul1, sul2, strA/B tet(A), TEM-1B</i>	<i>dfrA5</i>	–	–	IncFII, IncFIB, pBS512, pSL483
EC247_ST69H *	<i>dfrA7, sul1, sul2, strA/B, tet(A), TEM-1B</i>	<i>dfrA7</i>	–	–	IncFII, IncFIA, IncFIB, pCol4401
EC275_ST131C *	<i>aac-3-lll, aadA5, dfrA17, mph(A), sul1, sul2, strA/B, tet(A), TEM-1B</i>	<i>aadA5, dfrA17, sul1, mph(A)</i>		<i>gyrA(S83L)</i>	IncFII, IncFIB, Col156
EC398_ST4702C * \$	<i>aac-3-lll, aadA5, dfrA17, mph(A), sul1, TEM-1B</i>	<i>aadA5, dfrA17, sul1, mph(A)</i>	CTX-M-14	–	IncFII, IncFIB, IncFIA, Col156, pColBS512
EC297_ST131C * \$ #	<i>aadA5, dfrA17, mph(A), sul1, strA/B, sul2, tet(A)</i>	<i>aadA5, dfrA17, mph(A), sul1, strA/B, sul2</i>	CTX-M-27	<i>gyrA(S83L, D87N); parC(S80I, E84V)</i>	IncFII, IncFIB, IncFIA, IncB,
EC357_ST131C * \$ #	<i>aac-3-IIb, aacA4, dfrA14, strA/B, sul2, tet(A), TEM-1B</i>	<i>dfrA14</i>	CTX-M-15, OXA-1	<i>gyrA(S83L, D87N); parC(S80I, E84V)</i>	IncFII, IncFIB
EC231_ST38C * \$	<i>aadA5, dfrA17, mph(A), sul1, strA/B, sul2, tet(A)</i>	<i>aadA5, dfrA17, mph(A), sul1, strA/B, sul2</i>	CTX-M-27	–	IncFII, IncFIB
EC46_ST131H * \$ #	<i>aac-3-lla, aac(6')-1b-cr, aadA5, dfrA17, mph(A), sul1, sul2, strA, floR, TEM-1C</i>	<i>aadA5, dfrA17, mph(A), sul1, strA/B, sul2</i>	CTX-M-15, OXA-1	<i>gyrA(S83L, D87N); parC(S80I, E84V), aac(6')-1b-cr</i>	IncFII, IncFIB, IncFIA, IncX4, Col156
EC36_ST38H * \$ #	<i>aac-3-lll, aadA5, dfrA17, mph(A), sul1, sul2, strA/B, tet(A), TEM-1B</i>	<i>aadA5, dfrA17, mph(A), sul1, strA/B, sul2</i>	CTX-M-15	<i>gyrA(S83L, D87N); parC(S80I)</i>	IncFII, IncFIB, IncI1, Col156
EC169_ST73C * #	<i>aadA1, sul1, TEM-1B</i>	<i>aadA1, sul1</i>	–	–	IncFII, IncFIB, Col156
EC12_ST131H * #	<i>aadA5, dfrA17, mph(A), sul1, TEM-1B</i>	<i>aadA5, dfrA17, sul1, mph(A)</i>	–	<i>gyrA(S83L, D87N); parC(S80I, E84V)</i>	IncFII, IncFIA, IncI1, Col156
EC222_ST73H *	<i>aadA1, sul1, TEM-1B</i>	<i>aadA1</i>	–	–	IncFII, IncFIB, Col156
EC182_ST73C *	<i>aadA1, dfrA1, sul1, strA/B, sul2, TEM-1B</i>	<i>aadA1, dfrA1</i>	–	–	IncFII, IncFIB, IncFIA
EC149_ST73C *	<i>aadA1, dfrA14, strA/B, sul2, sul1, TEM-1B</i>	<i>aadA1, dfrA14</i>	–	–	IncFII, IncFIB, Col156

Continued



Table 1. Continued

Isolate descriptor	Resistance genes in the genome	Class 1 integron-associated genes	ESBL genes	Fluoroquinolone resistance genes	Plasmid replicons
EC8_ST131C *	<i>aadA5, dfrA17, mph(A), sul1, TEM-1B</i>	<i>aadA5, dfrA17</i>		–	IncFII, IncFIB, IncFIA, IncX4, Col156
EC233_ST8196C * \$ #	<i>qnrB4, aac-3-lla, aadA5, dfrA17, mph(A), sul1, blaDHA-1</i>	<i>aadA5, dfrA17, sul1, mph(A)</i>	CTX-M-15	<i>gyrA(S83L, D87N); parC(S80I, E84V), qnrB4</i>	IncFII, IncI1, Col156, pColBS512
EC120_ST131C * \$ #	<i>aac-3-Ib, aacA4, dfrA14, strA/B, sul2, TEM-1B</i>	<i>dfrA14</i>	CTX-M-15, OXA-1	<i>gyrA(S83L, D87N); parC(S80I, E84V)</i>	IncFII, IncFIB, IncFIA, IncI, pCol440II
EC44_ST131H * \$ #	<i>aac-3-lla, aac(6')-1b-cr, aadA5, dfrA17, mph(A), sul1, sul2, strA, floR, TEM-1C</i>	<i>aadA5, dfrA17, mph(A), sul1, strA/B, sul2</i>	CTX-M-15, OXA-1	<i>gyrA(S83L, D87N); parC(S80I, E84V), parE(I529L), aac(6')-1b-cr</i>	IncFII, IncFIB, IncFIA, IncX4, Col156
EC66_ST393C * \$ #	<i>aadA5, dfrA17, mph(A), sul1, ermB</i>	<i>aadA5, dfrA17, sul1, mph(A)</i>	CTX-M-15	<i>gyrA(S83L, D87N); parC(S57T, S80I), parE(L416F)</i>	IncFIB, IncFIA, Col156
EC87_ST131C * \$ #	<i>aadA5, dfrA17, mph(A), sul1, strA/B, sul2, tet(A)</i>	<i>aadA5, dfrA17, mph(A), sul1, strA/B, sul2</i>	CTX-M-27	<i>gyrA(S83L, D87N); parC(S80I, E84V)</i>	IncFII, IncFIA, pColBS512
EC164_ST73C *	<i>TEM-1B</i>	–	–	–	0
EC226_ST349C *	–	–	–	–	0
EC285_ST648C * #	–	–	–	<i>gyrA(S83L, D87N); parC(S80I)</i>	0
EC274_ST38H * #	<i>aac-3-lll, catA1, tet(B), TEM-1B</i>	–	–	<i>gyrA(S83L, D87N); parC(S80I, E84G)</i>	IncFII, IncFIB
EC282_ST73C *	<i>TEM-1B</i>	–	–	–	IncFII
EC179_ST131C * \$ #	–	–	CTX-M-27	<i>gyrA(S83L, D87N); parC(S80I, E84V)</i>	IncFII, IncFIB, IncFIA, IncB, Col156
EC168_ST73C *	<i>TEM-1B</i>	–	–	–	IncFII, IncFIB, IncX4, Col156
EC25_ST88H *	<i>aadA1, dfrA1, sul1, TEM-1B</i>	<i>aadA1, dfrA1</i>		–	0
EC393_ST131C * #	<i>TEM-1B</i>	–	–	<i>gyrA(S83L)</i>	0
EC302_ST349C	<i>dfrA17, sul1, sul2, strA/B, tet(A), TEM-1B</i>	<i>dfrA17</i>		–	IncFII, IncQ
EC272_ST73C	<i>aadA1, sul1, TEM-1B</i>	<i>aadA1</i>	–	–	IncFII, IncFIB, Col156
EC347_ST73C	<i>aadA1, sul1</i>	<i>aadA1</i>	SHV-48	–	IncFII, IncFIB, Col156
EC337_ST69H	<i>dfrA5, TEM-1B</i>	<i>dfrA5</i>		–	IncB, IncFIB
EC323_ST549H	<i>dfrA5, TEM-1B</i>	<i>dfrA5</i>	–	–	IncFII, IncFIB
EC330_ST681C	<i>dfrA5, sul1, strA/B, sul2, tet(A), TEM-1B</i>	<i>dfrA5</i>	–	–	IncB, ColBS512
EC290_ST127C	<i>dfrA5, sul1, TEM-1B</i>	<i>dfrA5</i>	–	–	IncFIB, Col156
EC273_ST1434C	<i>aph-3-la, dfrA5, sul2, strA/B, tet(A), TEM-1B</i>	<i>dfrA5</i>	–	–	IncFII, IncFIB, IncH11, IncX4
EC131_ST131C	<i>aadA5, dfrA17, mph(A), sul1, strA/B, sul2, tet(A)</i>	<i>aadA5, dfrA17</i>	–	<i>gyrA(S83L)</i>	IncFII, IncFIB, Col156
EC239_ST62C	<i>catA1, dfrA17</i>	<i>dfrA17</i>	–	–	IncFII, IncFIB, Col156

Continued

Table 1. Continued

Isolate descriptor	Resistance genes in the genome	Class 1 integron-associated genes	ESBL genes	Fluoroquinolone resistance genes	Plasmid replicons
EC147_ST73C	<i>aadA1, dfrA1, strA/B, sul1, sul2, TEM-1B</i>	<i>aadA1, dfrA1</i>	–	–	IncFII, IncFIB, IncI1, Col156
EC65_ST127C	<i>aadA1, sul1, TEM-1B</i>	<i>aadA1</i>	–	–	IncFII, IncFIB, Col156
EC173_ST12C	<i>aadA1, dfrA1, sul1, tet(B), TEM-1B</i>	<i>aadA1, dfrA1</i>	–	–	IncFIA, IncFIB, Col156, ColVC
EC108_95C	<i>aadA2, catA1, dfrA12, mph(A), sul1, tet(B), TEM-1B</i>	<i>dfrA12, aadA2</i>	–	–	IncFII, IncFIB, IncFIA, Col156
EC16_ST1193C #	<i>dfrA17, strA/B, sul2, tet(B), TEM-1B</i>	<i>dfrA17</i>	–	<i>gyrA(S83L, D87N); parC(S80I)</i>	IncFII, IncFIB, IncFIA, IncQ, Col156
EC5_ST73C	<i>aadA1, sul1, TEM-1B</i>	<i>aadA1</i>	–	–	IncFII, IncFIB, Col156
EC152_ST95C	<i>aadA1, aadA2, cmlA1, dfrA12, sul3</i>	<i>aadA1, orfE, aadA2, cmlA1, dfrA12, sul3</i>	–	–	IncFII, IncFIB
EC177_ST95H	<i>dfrA5</i>	<i>dfrA5</i>	–	–	IncFII, IncFIB, Col156
EC45_ST57C	<i>aadA1, dfrA5, sul1, sul2, tet(C)</i>	<i>aadA1, dfrA5</i>	–	–	IncFII, IncFIB
EC84_ST131H	<i>dfrA14, mphA</i>	<i>dfrA14</i>	–	–	IncFII, IncFIB, Col156
EC56_ST164C	–	–	–	–	0
EC2_ST95C	–	–	–	–	IncFIA, IncFIB, Col156
EC160_ST73C	–	–	–	–	IncFII, IncFIB, Col156
EC178_ST69C	<i>TEM-1B</i>	–	–	<i>gyrA(S83L)</i>	IncFII, IncFIB, IncFIA
EC250_ST127C	–	–	–	–	0
EC235_ST117C	–	–	–	–	IncFIB
EC230_ST91C	–	–	–	–	IncFII, IncFIB
EC280_ST144C	<i>TEM-1B</i>	–	OXA-4	–	IncFII, IncFIB, Col156
EC232_ST73C	–	–	–	–	0
EC249a_ST127H	–	–	–	–	IncFII, IncFIB
EC276_ST906C	–	–	–	–	IncFII, IncFIB
EC33_ST73C	<i>TEM-1B, sul2</i>	–	–	–	0
EC7_ST144C	<i>dfrA14, sul2, strA/B, TEM-1B</i>	–	–	–	IncB, IncFII, IncFIB, Col156
EC11_ST95C	<i>TEM-1B</i>	–	–	–	IncFII, IncFIB
EC51_ST69C	<i>sul1, TEM-1B</i>	–	–	–	IncB, IncFII, IncFIB, Col156
EC128_ST8197H	<i>TEM-1C</i>	–	–	–	IncFII, IncFIB, IncB, Col156
EC15_ST5242C	–	–	–	–	0

Continued

Table 1. Continued

Isolate descriptor	Resistance genes in the genome	Class 1 integron-associated genes	ESBL genes	Fluoroquinolone resistance genes	Plasmid replicons
EC19_ST372C	-	-	-	-	0
EC35_ST69H	-	-	-	-	IncB, IncFII, IncFIB, Col156
EC68_ST95C	<i>tet(B)</i>	-	-	-	IncFIA, IncFII, IncFIB
EC20_ST73C	-	-	-	-	0
EC372_ST95C	-	-	-	-	IncFIA, IncFIB, Col156
EC104_ST95C	<i>strA, strB, tet(B)</i>	-	-	<i>gyrA(S83L)</i>	IncFII, IncFIB
EC123_ST88C	<i>strA, strB, sul2</i>	-	-	-	IncFII, IncFIB, IncQ
EC24_ST73C	-	-	-	-	0
EC223_ST62C	<i>aadA1, dfrA1, tet(A)</i>	-	-	<i>gyrA(S83L)</i>	IncFIB, Col156

The isolate descriptor column consists of isolate number, the sequence type (ST) and hospital onset (H) or community onset (C) included as suffix, followed by symbols to include phenotypic profile, where MDR is indicated as \*, ESBL producers are indicated as \$ and fluoroquinolone-resistant isolates indicated as #.

positive for *kpsMT II*, while only isolate EC323 tested positive for *kpsMT III*. Fourteen isolates were positive for *ompT*, 58 for *traT*, 67 for *iss* and 11 for *cavC* (File S3). None of the isolates tested positive for *hlyA*, but 22 (all phylogroup B2) were positive for *cnf1*. No obvious trend was observed in the relative distribution of genes encoding known outer-membrane iron receptors, such as *fyuA*, *ireA*, *iroN* and *iutA*, amongst the different *E. coli* sequence types, phylogroups (File S4) or CO vs HO infections.

### Plasmid replicon diversity within the cohort

Plasmid typing identified IncFII and IncFIB replicons in most isolates (51/81), including 39 CO BSI and 12 HO BSI isolates (Table 1). In 11 isolates from CO infections and 3 from HO infections, FIA replicons were co-resident with IncFII and IncFIB replicons. We also identified small col-like plasmid replicons in 33 of the 51 isolates that carried IncFII and IncFIB replicons, raising the possibility of presence of large IncFII–IncFIB hybrid ColV-like virulence plasmids in them.

Two completely closed IncFII–IncFIB hybrid ColV-like plasmids, pCERC4 (Fig. S1a) and pSDJ2009-52F (Fig. S1b), have recently been reported and characterized [54, 55] from Sydney. Illumina read mapping of all 81 genomes against both plasmids identified 12 isolates which appeared to carry most of the virulence and colicin resistance genes characteristic of ColV hybrid virulence plasmids (Fig. S1a, b). To examine this further, a series of BLASTn analyses of the ColV operon, *col-Ia* genes, the virulence gene repertoire and their linkage to *repA* and *repA1* genes found on pCREC4 was conducted against all assembled genomes in our cohort. Combining data on scaffolds identified in our BLASTn analyses (File S5) and following the stringent parameters indicated in the Methods section,

our assessment identified genetic features that define ColV-like IncFII–IncFIB plasmids in 10 isolates (Table 2). Of these, five typed as ST95 (EC104, EC68, EC177, EC152, EC11), two each as ST131 (EC120 and EC357) and ST88 (EC25, EC123), and one as ST540 (EC236). Two genomes (EC337 and EC323) that had the Col-V operon and *col-Ia* genes did not contain both (IncFII and IncFIB) replicons characteristic of hybrid IncFII–IncFIB plasmids and likely represent isolates with novel ColV-like IncF plasmid variants. All 10 isolates identified in the analysis as candidates likely hosting IncFII–IncFIB ColV-like plasmids did not return positive hits for both ColV operon and *col-Ia* genes (Table 2), indicating subtle variations in the structure of plasmids in this group. Further studies with completely closed plasmid sequences are required to characterize them and examine the distribution of all ColV-like plasmid variants within isolates from the clinic and elsewhere.

We also identified an IncI1 replicon in six isolates (four associated with HO and two with CO infections), four of which had an *ISEcp1* associated with *bla*<sub>CTX-M-15</sub> (Table 1) gene. An IncX4 replicon was identified in five isolates, three of which were ST131. Finally, an IncQ replicon was identified in three MDR isolates belonging to ST38, ST349 and ST88.

## DISCUSSION

This study and others show that BSIs are caused by genetically diverse *E. coli* [56], indicating that pathogen virulence profiles and perhaps host immune status are likely to be important variables in disease causation. *E. coli* ST131, often isolated from patients with BSI, is widely disseminated in Europe, Asia, North America, Africa and Australia [32, 57], and has played a critical role in the global spread of *bla*<sub>CTX-M-15</sub>



**Fig. 3.** A presence/absence heatmap of *E. coli* virulence genes. The phylogenetic relationship of the isolates is presented as a cladogram along the y-axis and the virulence genes are listed in the top panel along the x-axis. Grey boxes indicate the presence and black boxes indicate the absence of a gene. Virulence gene profiling was not performed for the two reference genomes (EC958 and *E. coli* K12), hence the corresponding rows are left blank.

[58]. Our data show that *bla*<sub>CTX-M-15</sub> in BSIs in Sydney is not restricted to ST131 but extends to *E. coli* ST8196, ST393 and ST38. Chromosomal *bla*<sub>CTX-M-15</sub> in ST131 is usually associated with *ISEcpI* elements, but in EC44\_ST131, EC46\_ST131 and EC66\_ST393, *bla*<sub>CTX-M-15</sub> is not linked to *ISEcpI*, suggesting a different genetic context, including an association with other plasmids.

Although our data are based on a limited number of isolates acquired from a single hospital in Sydney, they highlight

similar population dynamics in the context of resistance and virulence genes found in invasive isolates reported on a cohort of more than 1500 *E. coli* collected over 11 years in the UK [57]. The number of virulence genes carried by ST131 was - less than that in the cohort of *E. coli* ST73 (Fig. 3, Files S3 and S4), suggesting that there is segregation between excessively virulent and extremely drug-resistant clonal lineages [59]. However, acquisition of virulence plasmids such as ColV that carry CRLs may alter this scenario, necessitating constant

**Table 2.** Summary of the BLASTn analysis of genomes with *colV* operon, *col-Ia* genes and the repertoire of virulence genes and the *repA* gene present on pCERC-4 (accession no. KU578032).

Isolate identifier	Sequence type	pCERC-4 <i>rep</i> genes	Positive hits to pCERC-4 virulence gene (n=20)
<b>Isolates with both <i>colV</i> operon and <i>col-Ia</i> genes</b>			
EC337_ST69H	69	<i>repA_FIB</i>	<i>iroC,iutA,iroN,etsB,iucA,iucC,etsC,iucD,iroD,etsA,iroB,hlyF,iroE,ompT,iucB,sitA,sitD,sitC,sitB,iss</i>
EC25_ST88H*	88	<i>repA_FIB</i> , <i>repA1_FII</i>	<i>iroC,iutA,iroN,etsA, etsB,etsC, iucA,iucC,iucD,iroD,iroB,hlyF,iroE,ompT,iucB,sitA,sitD,sitC,sitB,iss</i>
EC104_ST95C*	95	<i>repA_FIB</i> , <i>repA1_FII</i> <sup>§</sup>	<b><i>iroC,iutA,iroN,etsA, etsB,etsC,iucA,iucC,iucD,iroD,iroB,hlyF,iroE,ompT,iucB,sitC,sitB,sitA,sitD,iss</i></b>
EC68_ST95C*	95	<i>repA_FIB</i> , <i>repA1_FII</i> <sup>§</sup>	<i>iroC,iutA,iroN,etsA, etsB,etsC,iucA,iucC,iucD,iroD,iroB,hlyF,iroE,ompT,iucB,sitC,sitD,sitB,sitA,iss</i>
EC177_ST95H*	95	<i>repA_FIB</i> , <i>repA1_FII</i>	<i>iroC,iroN,etsA,etsB,etsC,iroD, iroB,hlyF,iroE,ompT,sitA,sitC,sitD,sitB,iss</i>
EC120_ST131C*	131	<i>repA_FIB</i> , <i>repA1_FII</i> <sup>§</sup>	<i>iroC,iutA,iroN,etsA, etsB,etsC,iucA,iucD,iroD,iroB,hlyF,iroE,ompT,iucB,sitC,sitB,iucC,sitA,sitD,iss</i>
EC236_ST540C*	540	<i>repA_FIB</i> , <i>repA1_FII</i>	<i>iroC,iutA,iroN,etsB,iucA,iucC,etsC,iucD,iroD,etsA,iroB,hlyF,iroE,ompT,iucB,sitA,sitD,sitC,sitB,iss</i>
EC323_ST549H*	549	<i>repA_FIB</i>	<b><i>iroC,iutA,iroN,etsA, etsB,etsC,iucA,iucC,iucD,iroD,iroB,hlyF,iroE,ompT,iucB,sitA,sitD,sitC,sitB,iss</i></b>
<b>Isolates with <i>colV</i> operon only</b>			
EC123_ST88C*	88	<i>repA_FIB</i> , <i>repA1_FII</i> <sup>§</sup>	<i>iroC,iutA,iroN,etsA, etsB,etsC,iucA,iucC,iucD,iroD,iroB,hlyF,iroE,ompT,iucB,sitA,sitD,sitC,sitB,iss</i>
EC152_ST95C*	95	<i>repA_FIB</i> , <i>repA1_FII</i> <sup>§</sup>	<i>iroC,iutA,iroN,etsA, etsB,etsC,iucA,iucC,iucD,iroD,iroB,hlyF,iroE,ompT,iucB,sitC,sitB,sitD,sitA,iss</i>
EC11_ST95C*	95	<i>repA_FIB</i> , <i>repA1_FII</i> <sup>§</sup>	<i>iroC,iutA,iroN,etsA, etsB,etsC,iucA,iucC,iucD,iroD,iroB,hlyF,iroE,ompT,iucB,sitC,sitB,sitA,sitD,iss</i>
EC357_ST131C*	131	<i>repA_FIB</i> , <i>repA1_FII</i> <sup>§</sup>	<i>iroC,iutA,iroN,etsA, etsB,etsC,iucA,iucD,iroD,iroB,hlyF,iucC,iroE,ompT,iucB,sitC,sitB,sitA</i>
<b>Isolates with <i>col-Ia</i> genes only</b>			
EC232_ST73C	73	None detected	<i>iroC,iutA,iroN,iucA,iucC,iucD,iroD,iroB,iroE,iucB,sitA,sitC,sitD,sitB,iss</i>
EC230_ST91C	91	<i>repA_FIB</i> <sup>§</sup> , <i>repA1_FII</i> <sup>§</sup>	<i>sitA,sitC,sitD,sitB,iss</i>
EC12_ST131H	131	<i>repA1_FII</i> <sup>§</sup>	<i>iutA,iucA,iucC,iucD,iucB,sitA,sitC,sitD,sitB,iss</i>

Asterisk (\*) on 'isolate identifiers' indicates linkage of *repA* with virulence genes on the same genomic scaffold (detailed in File S5). Virulence genes highlighted in bold fonts indicate genes that were assembled on a single scaffold with the *repA* gene, indicating linkage. A dollar (\$) sign in the pCERC-4 *rep* gene column represents  $\geq 95\%$  but not 100% identity with the respective *rep* gene sequence.

surveillance to identify emerging resistance threats [52]. While lineages of ST73 with MDR genotypes are commonly described, drug-sensitive lineages and lineages devoid of plasmids in Australia are also documented [60].

IncFII and IncFIB plasmid replicons were abundant in genetically diverse isolates acquired from both CO and HO BSIs. Virulence gene profiling studies suggested that a subset of *E. coli* isolates with ST95, ST131 and others likely carry an IncFII–IncFIB plasmid. Plasmids that have the ColV operon associated with IncF transfer regions often harbour a combination of virulence genes linked to serious human and animal diseases [24, 54, 61–65]. Based on an animal model of UPEC disease, it was demonstrated that acquisition of ColV-like

plasmid and associated virulence genes influence disease aetiology [66, 67]. In addition, the acquisition of a ColV-like plasmid by *Salmonella enterica* serovar Kentucky has been purported to be a key genetic event that transformed its ability to cause serious disease in poultry by conferring new colonization and fitness capabilities [24]. ColV plasmids are implicated in invasive avian diseases, such as colibacillosis caused by avian pathogenic *E. coli* (APEC) in commercial poultry, and are considered to be a defining trait of the ExPEC sub-pathotype [13, 61]. These plasmids carry an extensive suite of VAGSs necessary for colonization and disease formation in diverse hosts, including humans. Gene products that mediate increased serum survival (*iss*), membrane vesicle formation

protein (encoded by *hlyF* gene), outer-membrane protease activity (encoded by *ompT*), aerobactin biosynthesis (*iuc* operon), temperature-sensitive haemagglutination (encoded by *tsh*) and various (*sit*, *iroN* and *eit*) iron acquisition systems [5, 55] are significant in this regard. Our BLASTn analysis with assembled genomes supported initial observations from mapping of raw Illumina reads against two hybrid IncFII–IncFIB–ColV plasmids, indicating the presence of genes in the ColV operon in EC11\_ST95, EC152\_ST95C, EC123\_ST88C and EC357\_ST131C. Strains carrying the ColV operon and *col-Ia* genes included EC104\_ST95C, EC177\_ST95H, EC120\_ST131C, EC337\_ST69H, EC236\_ST540C and EC25\_ST88H. Our data also indicate that there are variabilities in the IncF plasmids circulating in our BSI cohort [54] and some may carry different combinations of the *colV* operon, *col-Ia* genes and a repertoire of virulence genes- the three features that define typical hybrid ColV plasmids. Further studies using long-read sequencing methodologies are required to address the diversity of plasmids carrying ColV-like virulence gene cargo and examine their associations with antibiotic resistance genes.

Forty-seven isolates in the bacteraemia cohort (58%) carry a class 1 integron (Table 1). As expected, most of the *intI1*-positive isolates carried a larger number of antibiotic resistance genes compared to isolates that did not have a class 1 integron. Almost all *E. coli* with *intI1* carried a minimum set of genes encoding resistance to streptomycin (*aadA* gene cassettes) and/or trimethoprim (*dfrA* gene cassettes), sulphonamide drugs (*sul1*, *sul2* or *sul3* genes) and ampicillin (*bla<sub>TEM-1</sub>*). Most isolates (34/49; 73%) that displayed phenotypic resistances to trimethoprim, used commonly as a first-line treatment option for uncomplicated UTIs, also had a sulphonamide resistance gene associated with the class 1 integron in the CRL. The class 1 integron is inarguably one of the most successful genetic elements responsible for the rapid evolution of multiple drug resistance, as it has captured a wide array of resistance gene cassettes [68] and dispersed within diverse animal and plant commensals and pathogens in every continent [69]. The structural boundaries that define class 1 integrons continue to evolve, often mediated by insertion of other mobile genetic elements such as IS26 [70–74] that is well documented to mobilize diverse antibiotic resistance genes. Class 1 integrons are found on various plasmid backbones, including conjugative broad-host-range plasmids, and is widely considered to be a reliable proxy for identification of multiple drug-resistant genomes [16, 75].

The MDR isolates in this cohort exhibited little diversity in the carriage of cassette-associated genes. Most class 1 integrons appeared to be variants of In2 or In4, some of which were associated with Tn21/Tn1721 hybrid transposon backbones. Our data identified IS26 adjacent to resistance genes in most of the CRLs (Fig. 3) and their presence creates scaffold breaks in the assembly of genome sequences. We also provide evidence for multiple IS26 elements per *E. coli* genome (Table 1), reinforcing the role played by IS26 in driving the evolution of multidrug resistant phenotypes in the Enterobacterales, including those causing BSIs. This is a

critically important observation because: (i) IS26 can mobilize a wide range of antibiotic resistance genes; (ii) IS26 is a preferred site for integration of translocatable units (TUs); and (iii) CRLs flanked by IS26 have the capacity to move laterally as independent TUs [76]. These observations emphasize the necessity for targeting mobile resistance scaffolds in developing and implementing comprehensive surveillance protocols for resistance gene epidemiology. Integration of such real-time surveillance data on the mobile resistome of pathogen and commensal populations in healthcare settings will better assist infection control teams to identify hotspots for the evolution of MDR pathogens.

One of the most significant findings from this study is the identification of ST8196 within the ST131 clonal complex. ST8196 isolates shared similar genotypic, phenotypic and plasmid profiles with ST131 isolates in the cohort. Both ST8196 isolates were identified from the 2014 sample pool and displayed resistance to multiple antibiotics, carriage of *bla<sub>CTX-M-15</sub>* and chromosomal *parC* and *gyrB* mutations known to confer fluoroquinolone resistance (Table 1). In addition, the ST8196 isolates have acquired a plasmid-associated *qnrB4* gene. The isolation of ST8196 with plasmid-encoded *qnrB4* genes indicates that active surveillance for this ST is needed in Sydney and Australia more broadly.

In conclusion, MDR isolates in this genetically diverse BSI cohort exhibit limited variation in class 1 integron structure and in the cassette arrays they carry. IS26 is playing a significant role in shaping the composition of genes in the CRLs, which mostly include class 1 integrons in pathogenic *E. coli*. The role played by IS26 in the evolution of CRL is significant because they present further opportunities to rapidly acquire antibiotic resistance genes, including ones encoding resistance to clinically important antibiotics. We have identified *E. coli* in this cohort of isolates from BSIs that carry the plasmid-associated fluoroquinolone resistance gene *qnrB4* and *bla<sub>CTX-M-15</sub>*. Dissemination of such plasmids may contribute to the rapid rise of co-resistance to two clinically important groups of antibiotics. Finally, IncF plasmids are important in the evolution of resistance and virulence attributes in *E. coli* causing BSIs in Australia. A detailed investigation of the plasmids that circulate in *E. coli* causing bacteraemia is warranted.

#### Funding information

This project was funded by the Australian Centre for Genomic Epidemiological Microbiology (AusGEM), a strategic partnership between the NSW Department of Primary Industries and the University of Technology Sydney. P. H. is the recipient of the University of Technology Sydney PhD scholarship as a part of the AusGEM programme and the University of Technology International Research Scholarship (IRS). M. C. is a recipient of an Australian Government Research Training Program Scholarship. These results are helping to inform the OUTBREAK decision support system funded by the Medical Research Future Fund Frontier Health and Medical Research Program.

#### Acknowledgements

We wish to acknowledge Michael Liu for assistance with NextEra Sequencing Library preparation, and Cameron Reid for helping with the bioinformatics workflow used in this manuscript.

**Author contributions**

P. H. was involved in data curation, formal analysis, investigation, worked on wet and dry lab methods described in the analysis, visualized the data and wrote a preliminary draft of the manuscript. M. C. was involved in the development of bioinformatics pipelines for ARIBA and plasmid read mapping analysis. He also generated some bioinformatics datasets for the project and assisted P. H. in setting up bioinformatics protocols used in this project. T. G. and E.C. were involved in project conceptualization, supplied strains analysed in the study and edited the manuscript. J. M. was involved in the generation of metadata supplied by the hospital and in data curation at the hospital. G. M. was involved project management and investigation, student supervision and edited the manuscript. S. P. D. was involved in conceptualization of the project, project management, supervision and reviewing the drafts of the manuscript and securing funds to carry out the research. P. R. C. was involved in conceptualization of the project, data curation and validation, formal analysis, investigation, methodology, visualization, writing the final versions of the manuscript, acquisition of AusGEM funding, project administration, management and supervision.

**Conflicts of interest**

The authors declare that there are no conflicts of interest.

**Ethical statement**

This study has been approved by the Sydney Local Health District CRGH Human Research Ethics committee as a 'Low and Negotiable Risk project' (approval ID: CH62/6/2016-093 – P Chowdhury LNR/16/CRGH/120), as we have not accessed hospital ID numbers or any other data that could potentially lead to the identification of individual patients. The isolates used in this retrospective study were from samples stored in the freezer at the Microbiology Laboratory at Concord Hospital. Therefore, we did not require any consent from patients.

**Data Bibliography**

GenBank accession number for reference genomes used in this manuscript were U00096.2 (*E. coli* K12), NZ\_HG941718.1 (*E. coli* EC958), CP000647.1 (*Klebsiella pneumoniae* subsp. *pneumoniae* MGH78578) and plasmid sequence ID KU578032 (pCERC-4) and MH195200.1 (pSDJ2009-52F).

**References**

- Diekema DJ, Hsueh P-R, Mendes RE, Pfaller MA, Rolston KV *et al*. The microbiology of bloodstream infection: 20-year trends from the SENTRY antimicrobial surveillance program. *Antimicrob Agents Chemother* 2019;63.
- Institute TG. *Stopping Sepsis A National Action Plan*. Sydney, Australia: Australian Sepsis Network; 2017.
- Coombs G, Bell JM, Daley D, Collingnon P, Cooley L *et al*. Australian group on antimicrobial resistance sepsis outcomes programs: 2017 report. report. Sydney2019. Report No.: Sydney: ACSQHC 2019.
- Turnidge J, Gottlieb T, Bell JM. Enterobacteriaceae sepsis outcome programme (EnSOP) 2013 antimicrobial susceptibility report 2014.
- Ananias M, Yano T. Serogroups and virulence genotypes of *Escherichia coli* isolated from patients with sepsis. *Braz J Med Biol Res* 2008;41:877–883.
- Johnson JR. Virulence factors in *Escherichia coli* urinary tract infection. *Clin Microbiol Rev* 1991;4:80–128.
- Coombs G, Bell JM, Daley D, Collingnon P, Cooley L *et al*. Australian group on antimicrobial resistance. sepsis outcome programs 2016 report. *Sydney2018 Contract No.: Sydney: ACSQHC* 2018.
- Abernethy J, Guy R, Sheridan EA, Hopkins S, Kiernan M *et al*. Epidemiology of *Escherichia coli* bacteraemia in England: results of an enhanced sentinel surveillance programme. *J Hosp Infect* 2017;95:365–375.
- Gillings M, Boucher Y, Labbate M, Holmes A, Krishnan S *et al*. The evolution of class 1 integrons and the rise of antibiotic resistance. *J Bacteriol* 2008;190:5095–5100.
- Hall RM, Stokes HW. Integrons: novel DNA elements which capture genes by site-specific recombination. *Genetica* 1993;90:115–132.
- Gündoğdu A, Long YB, Vollmerhausen TL, Katouli M. Antimicrobial resistance and distribution of sul genes and integron-associated intl genes among uropathogenic *Escherichia coli* in Queensland, Australia. *J Med Microbiol* 2011;60:1633–1642.
- Leverstein-van Hall MA, M Blok HE, T Donders AR, Paauw A, Fluit AC *et al*. Multidrug resistance among *Enterobacteriaceae* is strongly associated with the presence of integrons and is independent of species or isolate origin. *J Infect Dis* 2003;187:251–259.
- Cummins ML, Reid CJ, Roy Chowdhury P, Bushell RN, Esbert N *et al*. Whole genome sequence analysis of Australian avian pathogenic *Escherichia coli* that carry the class 1 integrase gene. *Microb Genom* 2019;5.
- Djordjevic SP, Stokes HW, Roy Chowdhury P. Mobile elements, zoonotic pathogens and commensal bacteria: conduits for the delivery of resistance genes into humans, production animals and soil microbiota. *Front Microbiol* 2013;4:86.
- Reid CJ, Wyrsh ER, Roy Chowdhury P, Zingali T, Liu M *et al*. Porcine commensal *Escherichia coli*: a reservoir for class 1 integrons associated with IS26. *Microb Genom* 2017;3.
- Gillings MR. Integrons: past, present, and future. *Microbiol Mol Biol Rev* 2014;78:257–277.
- Copur-Cicek A, Ozgumus OB, Saral A, Sandalli C. Antimicrobial resistance patterns and integron carriage of *Escherichia coli* isolates causing community-acquired infections in Turkey. *Ann Lab Med* 2014;34:139–144.
- Horcajada JP, Soto S, Gajewski A, Smithson A, Jiménez de Anta MT *et al*. Quinolone-resistant uropathogenic *Escherichia coli* strains from phylogenetic group B2 have fewer virulence factors than their susceptible counterparts. *J Clin Microbiol* 2005;43:2962–2964.
- Johnson JR, Kuskowski MA, O'Bryan TT, Maslow JN. Epidemiological correlates of virulence genotype and phylogenetic background among *Escherichia coli* blood isolates from adults with diverse-source bacteremia. *J Infect Dis* 2002;185:1439–1447.
- Johnson JR, Kuskowski MA, Owens K, Gajewski A, Winokur PL. Phylogenetic origin and virulence genotype in relation to resistance to fluoroquinolones and/or extended-spectrum cephalosporins and cephamycins among *Escherichia coli* isolates from animals and humans. *J Infect Dis* 2003;188:759–768.
- Moreno E, Prats G, Sabaté M, Pérez T, Johnson JR *et al*. Quinolone, fluoroquinolone and trimethoprim/sulfamethoxazole resistance in relation to virulence determinants and phylogenetic background among uropathogenic *Escherichia coli*. *J Antimicrob Chemother* 2006;57:204–211.
- Carattoli A. Plasmids and the spread of resistance. *Int J Med Microbiol* 2013;303:298–304.
- Rozwandowicz M, Brouwer MSM, Fischer J, Wagenaar JA, Gonzalez-Zorn B *et al*. Plasmids carrying antimicrobial resistance genes in Enterobacteriaceae. *J Antimicrob Chemother* 2018;73:1121–1137.
- Crossman LC, Chaudhuri RR, Beatson SA, Wells TJ, Desvaux M *et al*. A commensal gene bad: complete genome sequence of the prototypical enterotoxigenic *Escherichia coli* strain H10407. *J Bacteriol* 2010;192:5822–5831.
- Gaastra W, Svennerholm AM. Colonization factors of human enterotoxigenic *Escherichia coli* (ETEC). *Trends Microbiol* 1996;4:444–452.
- Pasqua M, Michelacci V, Di Martino ML, Tozzoli R, Grossi M *et al*. The Intriguing Evolutionary Journey of Enteroinvasive *E. coli* (EIEC) toward Pathogenicity. *Front Microbiol* 2017;8:2390.
- Croxall G, Hale J, Weston V, Manning G, Cheetham P *et al*. Molecular epidemiology of extraintestinal pathogenic *Escherichia coli* isolates from a regional cohort of elderly patients highlights the prevalence of ST131 strains with increased antimicrobial resistance in both community and hospital care settings. *J Antimicrob Chemother* 2011;66:2501–2508.
- Riley LW. Pandemic lineages of extraintestinal pathogenic *Escherichia coli*. *Clin Microbiol Infect* 2014;20:380–390.
- Andersen PS, Stegger M, Aziz M, Contente-Cuomo T, Gibbons HS *et al*. Complete Genome Sequence of the Epidemic and Highly

- Virulent CTX-M-15-Producing H30-Rx Subclone of *Escherichia coli* ST131. *Genome Announc* 2013;1:e00988-13.
30. Price LB, Johnson JR, Aziz M, Clabots C, Johnston B et al. The epidemic of extended-spectrum- $\beta$ -lactamase-producing *Escherichia coli* ST131 is driven by a single highly pathogenic subclone, H30-Rx. *mBio* 2013;4:e00377-00313.
  31. Petty NK, Ben Zakour NL, Stanton-Cook M, Skippington E, Totsika M et al. Global dissemination of a multidrug resistant *Escherichia coli* clone. *Proc Natl Acad Sci U S A* 2014;111:5694-5699.
  32. Totsika M, Beatson SA, Sarkar S, Phan M-D, Petty NK et al. Insights into a multidrug resistant *Escherichia coli* pathogen of the globally disseminated ST131 lineage: genome analysis and virulence mechanisms. *PLoS One* 2011;6:e26578.
  33. Poirel L, Gniadkowski M, Nordmann P. Biochemical analysis of the ceftazidime-hydrolyzing extended-spectrum beta-lactamase CTX-M-15 and of its structurally related beta-lactamase CTX-M-3. *J Antimicrob Chemother* 2002;50:1031-1034.
  34. Poirel L, Kämpfer P, Nordmann P. Chromosome-encoded Ambler class A beta-lactamase of *Kluyvera georgiana*, a probable progenitor of a subgroup of CTX-M extended-spectrum beta-lactamases. *Antimicrob Agents Chemother* 2002;46:4038-4040.
  35. Shaikh S, Fatima J, Shakil S, Rizvi SMD, Kamal MA. Antibiotic resistance and extended spectrum beta-lactamases: types, epidemiology and treatment. *Saudi J Biol Sci* 2015;22:90-101.
  36. Bevan ER, Jones AM, Hawkey PM. Global epidemiology of CTX-M  $\beta$ -lactamases: temporal and geographical shifts in genotype. *J Antimicrob Chemother* 2017;72:2145-2155.
  37. Hwang J-H, Shin G-W, Hwang J-H, Lee C-S. Bloodstream infection due to CTX-M-15 and TEM-1 extended-spectrum  $\beta$ -lactamase-producing *Salmonella enterica* serovar Virchow ST16. *Jpn J Infect Dis* 2017;70:308-310.
  38. Paterson DL, Bonomo RA. Extended-Spectrum beta-lactamases: a clinical update. *Clin Microbiol Rev* 2005;18:657-686.
  39. Paterson DL, Hujer KM, Hujer AM, Yeiser B, Bonomo MD et al. Extended-spectrum beta-lactamases in *Klebsiella pneumoniae* bloodstream isolates from seven countries: dominance and widespread prevalence of SHV- and CTX-M-type beta-lactamases. *Antimicrob Agents Chemother* 2003;47:3554-3560.
  40. Third Australian report on antimicrobial use and resistance in human health- AURA 2019-May. 2019. [database on the Internet] 2019. Available from. <https://www.safetyandquality.gov.au/our-work/antimicrobial-resistance/antimicrobial-use-and-resistance-australia-surveillance-system-aura/aura-2019>
  41. National Alert System for Critical Antimicrobial Resistances (CARAlert). 2018. [database on the Internet]. Available from. <https://www.safetyandquality.gov.au/our-work/antimicrobial-resistance/antimicrobial-use-and-resistance-australia-surveillance-system-aura/national-alert-system-critical-antimicrobial-resistances-caralert>
  42. Bell JM, Turnidge JD, Coombs GW, Daley DA, Gottlieb T et al. Australian group on antimicrobial resistance Australian Enterobacteriaceae sepsis outcome programme, annual report 2014. Australia: the Department of Health 2016 contract No.: 2..
  43. Magiorakos A-P, Srinivasan A, Carey RB, Carmeli Y, Falagas ME et al. Multidrug-Resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. *Clin Microbiol Infect* 2012;18:268-281.
  44. Darling AE, McKinnon J, Worden P, Santos J, Charles IG et al. A draft genome of *Escherichia coli* sequence type 127 strain 2009-46. *Gut Pathog* 2014;6:32.
  45. Coil D, Jospin G, Darling AE. A5-miseq: an updated pipeline to assemble microbial genomes from Illumina MiSeq data. *Bioinformatics* 2015;31:587-589.
  46. Brettin. RastTK. *Scientific Reports* 2015;8:365.
  47. Hunt M, Mather AE, Sánchez-Busó L, Page AJ, Parkhill J et al. ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads. *Microb Genom* 2017;3:e000131.
  48. Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O et al. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother* 2014;58:3895-3903.
  49. Darling AE, Jospin G, Lowe E, Matsen FA, Bik HM et al. PhyloSift: phylogenetic analysis of genomes and metagenomes. *PeerJ* 2014;2:e243.
  50. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer; 2016.
  51. Clermont O, Christenson JK, Denamur E, Gordon DM. The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. *Environ Microbiol Rep* 2013;5:58-65.
  52. Vangchhia B, Abraham S, Bell JM, Collignon P, Gibson JS et al. Phylogenetic diversity, antimicrobial susceptibility and virulence characteristics of phylogroup F *Escherichia coli* in Australia. *Microbiology* 2016;162:1904-1912.
  53. Santos ACM, Zidko ACM, Pignatari AC, Silva RM. Assessing the diversity of the virulence potential of *Escherichia coli* isolated from bacteremia in São Paulo, Brazil. *Braz J Med Biol Res* 2013;46:968-973.
  54. McKinnon J, Roy Chowdhury P, Djordjevic SP. Genomic analysis of multidrug-resistant *Escherichia coli* ST58 causing urosepsis. *Int J Antimicrob Agents* 2018;52:430-435.
  55. Moran RA, Hall RM. Evolution of regions containing antibiotic resistance genes in FII-2-FIB-1 ColV-Colla virulence plasmids. *Microb Drug Resist* 2018;24:411-421.
  56. Daga AP, Koga VL, Soncini JGM, de Matos CM, Perugini MRE et al. *Escherichia coli* bloodstream infections in patients at a university hospital: virulence factors and clinical characteristics. *Front Cell Infect Microbiol* 2019;9:191.
  57. Day MJ, Doumith M, Abernethy J, Hope R, Reynolds R et al. Population structure of *Escherichia coli* causing bacteraemia in the UK and Ireland between 2001 and 2010. *J Antimicrob Chemother* 2016;71:2139-2142.
  58. Coque TM, Novais A, Carattoli A, Poirel L, Pitout J et al. Dissemination of clonally related *Escherichia coli* strains expressing extended-spectrum beta-lactamase CTX-M-15. *Emerg Infect Dis* 2008;14:195-200.
  59. Kallonen T, Brodrick HJ, Harris SR, Corander J, Brown NM et al. Systematic longitudinal survey of invasive *Escherichia coli* in England demonstrates a stable population structure only transiently disturbed by the emergence of ST131. *Genome Res* 2017;27:1437-.
  60. Bogema DR, McKinnon J, Liu M, Hitchick N, Miller N et al. Whole-genome analysis of extraintestinal *Escherichia coli* sequence type 73 from a single hospital over a 2 year period identified different circulating clonal groups. *Microb Genom* 2019.
  61. Johnson TJ, Siek KE, Johnson SJ, Nolan LK. DNA sequence of a ColV plasmid and prevalence of selected plasmid-encoded virulence genes among avian *Escherichia coli* strains. *J Bacteriol* 2006;188:745-758.
  62. Johnson TJ, Wannemuehler Y, Johnson SJ, Stell AL, Doetkott C et al. Comparison of extraintestinal pathogenic *Escherichia coli* strains from human and avian sources reveals a mixed subset representing potential zoonotic pathogens. *Appl Environ Microbiol* 2008;74:7043-7050.
  63. Peigne C, Bidet P, Mahjoub-Messai F, Plainvert C, Barbe V et al. The plasmid of *Escherichia coli* strain S88 (O45:K1:H7) that causes neonatal meningitis is closely related to avian pathogenic *E. coli* plasmids and is associated with high-level bacteremia in a neonatal rat meningitis model. *Infect Immun* 2009;77:2272-2284.
  64. Ron EZ. Host specificity of septicemic *Escherichia coli*: human and avian pathogens. *Curr Opin Microbiol* 2006;9:28-32.
  65. Reid CJ, McKinnon J, Djordjevic SP. Clonal ST131-H22 *Escherichia coli* strains from a healthy pig and a human urinary tract infection carry highly similar resistance and virulence plasmids. *Microb Genom* 2019;5.



66. Skyberg JA, Johnson TJ, Johnson JR, Clabots C, Logue CM *et al.* Acquisition of avian pathogenic *Escherichia coli* plasmids by a commensal *E. coli* isolate enhances its abilities to kill chicken embryos, grow in human urine, and colonize the murine kidney. *Infect Immun* 2006;74:6287–6292.
67. Zhu Ge X, Jiang J, Pan Z, Hu L, Wang S *et al.* Comparative genomic analysis shows that avian pathogenic *Escherichia coli* isolate IMT5155 (O2:K1:H5; ST complex 95, ST140) shares close relationship with ST95 APEC O1:K1 and human ExPEC O18:K1 strains. *PLoS One* 2014;9:e112048.
68. Partridge SR, Tsafnat G, Coiera E, Iredell JR. Gene cassettes and cassette arrays in mobile resistance integrons. *FEMS Microbiol Rev* 2009;33:757–784.
69. An X-L, Chen Q-L, Zhu D, Zhu Y-G, Gillings MR *et al.* Impact of wastewater treatment on the prevalence of integrons and the genetic diversity of integron gene cassettes. *Appl Environ Microbiol* 2018;84 [Epub ahead of print 01 05 2018].
70. Reid CJ, Roy Chowdhury P, Djordjevic SP. Tn6026 and Tn6029 are found in complex resistance regions mobilised by diverse plasmids and chromosomal islands in multiple antibiotic resistant Enterobacteriaceae. *Plasmid* 2015;80:127–137.
71. Roy Chowdhury P, Charles IG, Djordjevic SP. A role for Tn6029 in the evolution of the complex antibiotic resistance gene loci in genomic island 3 in enteroaggregative hemorrhagic *Escherichia coli* O104:H4. *PLoS One* 2015;10:e0115781.
72. Roy Chowdhury P, Ingold A, Vanegas N, Martínez E, Merlino J *et al.* Dissemination of multiple drug resistance genes by class 1 integrons in *Klebsiella pneumoniae* isolates from four countries: a comparative study. *Antimicrob Agents Chemother* 2011;55:3140–3149.
73. Roy Chowdhury P, McKinnon J, Liu M, Djordjevic SP. Multidrug resistant uropathogenic *Escherichia coli* ST405 with a novel, composite IS26 transposon in a unique chromosomal location. *Front Microbiol* 2018;9:3212.
74. Dawes FE, Kuzevski A, Bettelheim KA, Hornitzky MA, Djordjevic SP *et al.* Distribution of class 1 integrons with IS26-mediated deletions in their 3'-conserved segments in *Escherichia coli* of human and animal origin. *PLoS One* 2010;5:e12754.
75. Gillings MR. Class 1 integrons as invasive species. *Curr Opin Microbiol* 2017;38:10–15.
76. Harmer CJ, Hall RM. IS26-mediated precise excision of the IS26-aphA1a translocatable unit. *mBio* 2015;6:e01866–01815.

### Five reasons to publish your next article with a Microbiology Society journal

1. The Microbiology Society is a not-for-profit organization.
2. We offer fast and rigorous peer review – average time to first decision is 4–6 weeks.
3. Our journals have a global readership with subscriptions held in research institutions around the world.
4. 80% of our authors rate our submission process as 'excellent' or 'very good'.
5. Your article will be published on an interactive journal platform with advanced metrics.

Find out more and submit your article at [microbiologyresearch.org](http://microbiologyresearch.org).