Science of the Total Environment 949 (2024) 175079

Contents lists available at ScienceDirect

Science of the Total Environment

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

Comparative genomic analysis of ESBL-selected and non-selected *Escherichia coli* in Australian wastewater: Elucidating differences in diversity, antimicrobial resistance, and virulence profiles

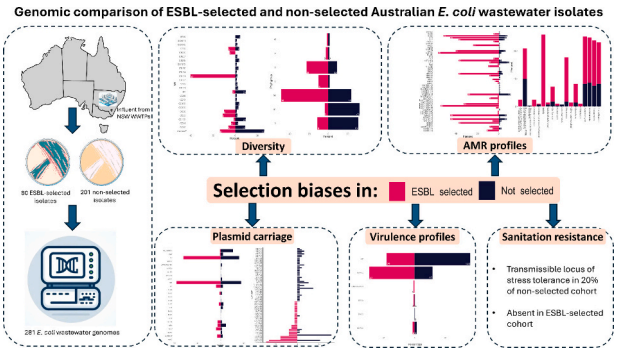
Veronica M. Jarocki [a,b,c,*], Dmitriy Li [a,b], Daniel R. Bogema [b,d], Jerald Yam [d], Cheryl Jenkins [b,d], Faisal I. Hai [e], Steven P. Djordjevic [a,b,*]

[a] Australian Institute for Microbiology & Infection, University of Technology Sydney, NSW, Australia
[b] Australian Centre for Genomic Epidemiological Microbiology, University of Technology Sydney, NSW, Australia
[c] Cooperative Research Centre for Solving Antimicrobial resistance in Agribusiness, Food and Environment (CRC SAAFE), South Australia, Australia
[d] Elizabeth Macarthur Agricultural Institute, NSW Department of Primary Industries, NSW, Australia
[e] School of Civil, Mining, Environmental and Architectural Engineering, University of Wollongong, NSW, Australia

HIGHLIGHTS

- ESBL-selected cohort showed reduced *Escherichia coli* diversity and an ST131 bias.
- High MDR prevalence in ESBL-selected isolates; no ARGs in 80 % of non-selected isolates.
- ExPEC were the dominate pathotype in both groups.
- Stress tolerance locus notable in non-selected isolates, absent in ESBL-selected group.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Kyle Bibby

Keywords:
Wastewater
WGS
AMR
tLST
ST131

ABSTRACT

Extended-spectrum β-lactamases (ESBLs)-producing *E. coli* have been proposed as an indicator bacterium for antimicrobial resistance (AMR) surveillance within a OneHealth framework. However, it is important to understand the effects and potential biases ESBL-selection has on *E. coli* populations. Utilising whole genome sequencing, this study compared 80 ESBL-selected *E. coli* isolates with 201 non-selected isolates from Australian wastewater. The findings revealed significant variations between these cohorts in genetic diversity, AMR profiles, and carriage of virulence-associated genes (VAGs), plasmids, and the transmissible Locus of Stress Tolerance (tLST), a genomic island that imparts resistance to extreme heat and chlorination. The study highlights the predominance of certain sequence types (STs), particularly ST131 (75 % clade A), in ESBL-selected isolates (40 % vs 2 %) and overall the ESBL-selected isolates were largely multidrug-resistant (MDR), predominantly carrying genes for resistance to aminoglycosides, extended-spectrum β-lactams, fluoroquinolone, macrolides, sulphonamides/trimethoprim, and tetracyclines. The ESBLs identified were almost exclusively $bla_{CTX\text{-}M}$ genes, most

* Corresponding authors at: Australian Institute for Microbiology & Infection, University of Technology Sydney, NSW, Australia.
E-mail addresses: veronica.jarocki@uts.edu.au (V.M. Jarocki), steven.djordjevic@uts.edu.au (S.P. Djordjevic).

https://doi.org/10.1016/j.scitotenv.2024.175079
Received 13 March 2024; Received in revised form 7 July 2024; Accepted 25 July 2024
Available online 31 July 2024

commonly $bla_{CTX\text{-}M\text{-}15}$ > $bla_{CTX\text{-}M\text{-}27}$ > $bla_{CTX\text{-}M\text{-}14}$. These were predominately carried on IncF plasmids or chromosomally (always IS*Ecp1* associated), in equal numbers. In contrast, 80 % of non-selected isolates carried no acquired ARGs, and none carried $bla_{CTX\text{-}M}$ genes. In both cohorts, extraintestinal pathogenic *E. coli* (ExPEC) was the dominate pathotype (35 % total) with few (4 % total) intestinal pathogenic *E. coli* pathotypes identified (aEPEC > ETEC > EAEC). Nevertheless, some clinically important genes were only identified in the non-selected group, namely tigecycline-resistance gene *tet(X4)* and AmpC ESBL $bla_{CMY\text{-}2}$. Additionally, the presence of tLST, associated with higher metal resistance gene carriage (Ag, As, Cu, Hg, Ni), in a substantial portion of non-selected isolates (20 % vs 0 %), underscores environmental pressures shaping bacterial populations in wastewater ecosystems. These insights are important for developing comprehensive, less biased genomic surveillance strategies to understand and manage public health threats posed by pathogenic *E. coli* and AMR.

1. Introduction

Antimicrobial resistance (AMR) poses a significant public and global health challenge, threatening health, food, and environmental safety. To tackle this issue, genomic surveillance programmes are being established to monitor the prevalence of resistant organisms (Baker et al., 2023; Djordjevic et al., 2023). Municipal wastewater treatment plants (WWTPs), which house bacterial populations from defined regions, have been proposed as potential AMR surveillance sites (Hendriksen et al., 2019). Monitoring antibiotic-resistant lineages, antimicrobial resistance genes (ARGs), and mobile genetic elements (MGEs) in untreated sewage (i.e., influent) could provide an indicator for regional AMR and its temporal changes. Wastewater environments, characterised by abundant nutrients and high microbial concentrations, provide an optimal setting for microbial growth and gene transfer between bacteria (Osińska et al., 2017). The ability of bacteria to transfer genes via MGEs, such as plasmids, phages, and transposons, plays a pivotal role in both pathogen evolution and the spread of AMR across microbial communities (Osińska et al., 2017). Wastewater is also composed of a broad spectrum of inorganic compounds, including antibiotics, detergents, personal care products, and metals. These compounds exert evolutionary selection pressures potentially driving survival strategies like treatment resistance and contribute to the evolution of naturalised populations (i.e., bacteria that have undergone genetic adaptations that enable them to thrive and persist in wastewater environments) (Zhi et al., 2019).

Escherichia coli primarily resides in the lower intestinal tracts of warm-blooded animals, and is discharged into the environment via faeces, urine, and wastewater. Historically, the presence of *E. coli* in the environment has served as an indicator of recent faecal contamination (Leclerc et al., 2001). However, research has demonstrated that *E. coli* can proliferate outside a host, adapting to diverse natural habitats, including soil and water, potentially complicating its role as a faecal indicator (Jang et al., 2017). *E. coli* can also persist in wastewater and recent investigations have revealed its ability to survive wastewater treatment processes (Behruznia and Gordon, 2022; Zhi et al., 2019, 2016). Most environmental *E. coli*, including wastewater-associated *E. coli* strains, have reportedly belonged to phylogroups A and B1. However, naturalised wastewater strains were found to form a distinct cluster within phylogroup A, specifically in two sequence types (STs), ST399 and ST635 (both belonging to clonal complex [CC] 399) (Zhi et al., 2019), indicating a potential evolutionary divergence from enteric *E. coli* strains (Yu et al., 2022). These naturalised *E. coli* strains frequently carry the transmissible Locus of Stress Tolerance (tLST) (also referred to as the Locus of Heat Resistance (LoHR)), a genomic region that confers an extreme heat-resistance phenotype along with greater resistance to chlorination and other advanced oxidants (Wang et al., 2020; Yu et al., 2022), all of which underpin microbial safety strategies. As such, these loci are of interest for public health, food safety and environmental microbiology, as they can enhance the survival and persistence of pathogenic bacteria in wastewater treatment plants (Yu et al., 2022). In addition to naturalised strains, certain *E. coli* pathotypes have also exhibited higher resistance to wastewater treatment processes. Studies have shown that strains surviving wastewater treatment were predominantly extraintestinal pathogenic *E. coli* (ExPEC) associated with urinary tract and bloodstream infections (Anastasi et al., 2013; Zhi et al., 2020), including pandemic extended-spectrum β-lactamase (ESBL)-producing ST131 strains (Dolejska et al., 2011; Zhi et al., 2020).

ESBL-producing bacteria are resistant to most β-lactam antibiotics, including 3rd generation cephalosporins, necessitating treatment with "last-resort" antibiotics. These bacteria are often multidrug resistance (MDR), particularly those listed as World Health Organisation (WHO) critical priority pathogens, including *E. coli*. The predominate ESBL genes currently circulating worldwide belong to the $bla_{CTX\text{-}M}$ family (> 200 variants), and in *E. coli* these genes are most commonly situated on IncF and IncI1 plasmids (Cormier et al., 2022; Pitout and Chen, 2023; Rozwandowicz et al., 2018). In recent years, the detection of ESBL-producing *E. coli* in wastewater has gained significant global attention due to its implications for public health and the environment (Conforti et al., 2024; Bréchet et al., 2014; Davidova-Gerzova et al., 2023). Globally, the prevalence of ESBL-producing *E. coli* in wastewater is influenced by local antibiotic usage patterns, industrial practices, and wastewater management practices. However, levels of ESBL-producing *E. coli* tend to be highest in hospital wastewater (Bréchet et al., 2014; Zaatout et al., 2021). The importance of monitoring ESBL-producing *E. coli* is encapsulated by the WHO and Advisory Group on Integrated Surveillance of Antimicrobial Resistance (AGISAR) developed Tricycle protocol, which describes a OneHealth AMR surveillance system using ESBL-producing *E. coli* as the singular indicator organism (WHO, 2021). However, concerns have emerged regarding the potential biases introduced by exclusively focusing on ESBL-producing strains in genomic surveillance efforts. The Tricycle protocol's reliance on ESBL-producing *E. coli* may inadvertently limit the comprehensive understanding of *E. coli* phylogenetic lineages and plasmids contributing to human and animal disease, particularly as certain ExPEC strains causing urinary tract infections (UTIs) often lack ESBLs (Djordjevic et al., 2023; Li et al., 2023). Relying solely on ESBL-producing *E. coli* for genomic surveillance may lead to overlooking emerging pathogens, reservoirs, and transmission vectors that fall outside the selected criteria (Djordjevic et al., 2023). Notably however, there is currently a paucity of genomic research directly comparing ESBL-producing *E. coli* with non-selected *E. coli* taken from the same sources, as are studies utilising whole genome sequencing (WGS) to investigate *E. coli* populations in wastewater in certain geographical regions, such as Australia.

We hypothesis that expanding genomic surveillance beyond ESBL-producing strains to include a broader range of *E. coli* types will provide a more comprehensive understanding of the epidemiology and evolution of this bacterium, particularly in environments such as municipal wastewater that are conducive to lateral gene transfer. We also aim to initiate efforts to address the current knowledge gap in genomic data on wastewater *E. coli* from Australia. Here we use WGS to compare the diversity of *E. coli* in Australian wastewater and analyse the distribution of ARGs, virulence associated genes (VAGs), metal resistance genes (MRGs), tLST, and plasmids across two distinct cohorts - one comprised of wastewater isolates selected for extended spectrum beta-lactam resistance ($n = 80$) and the other consisting of wastewater isolates not selected for any antibiotic phenotype ($n = 201$).

2. Materials and methods

2.1. Sample collection

Raw wastewater (influent) samples were collected from two WWTPs in New South Wales (NSW), Australia. WWTP1 serves an urban area of 71 km^2 and a population of approximately 200,000 while WWTP2 serves a smaller urban population of approximately 78,000. In 2020, a total of three 500 mL autosampler wastewater samples were collected: two samples from WWTP1, and one from WWTP2. The first sampling event for both WWPT1 and WWTP2 took place in the morning of 12th March 2020, between 08:00 and 09:00. A second sampling event for WWTP1 was conducted on 19th March 2020, at the same time. Due to COVID-19 pandemic restrictions in place at the time, it was not possible to conduct a second sampling at WWTP2. Regarding sampling locations, at WWTP2 the autosampler was positioned directly downstream of the step screen. Due to the presence of bolted-down covers at WWTP1, which restricted access to sampling points between screens, the autosampler was placed upstream of the step screens. All samples were collected in sterile containers and processed within 2–3 h of collection.

2.2. Sample processing

Upon receipt, the samples were processed to prepare non-enriched and buffered peptone water (BPW)-enriched cultures. For initial storage, 5 mL of the sample was aliquoted and combined with 40 % LB and glycerol (1:1), then stored at −80 °C.

For non-enriched cultures, serial dilution was performed by transferring 1 mL of stock sample into 9 mL of phosphate-buffered saline (PBS) and executing a 10-fold dilution twice, achieving final dilutions of 10^{-1} and 10^{-2}. Utilising a 10 μL blue loop, the neat stock sample and the two dilutions were spread onto both Chromogenic ECC Agar Plates [Edwards, Australia] and CHROMagarT ESBL Plates [Edwards, Australia] and incubated at 37 °C in an O_2 incubator for 16 to 18 h. Target colonies were then subcultured onto Nutrient Agar (NA) plates [ThermoFisher Scientific] and incubated overnight or for 16 to 18 h at 37 °C in an O_2 incubator. Target colonies from NA plates were transferred using a sterile toothpick onto a MALDI plate for species identification.

For BPW-enriched cultures, 1 mL of stock sample was added to 9 mL of BPW in a 30 mL tube. This mixture was incubated at 37 °C in an O_2 incubator for 16 to 18 h. Serial dilution of the BPW-enriched samples was then conducted by transferring 100 μL into 900 μL of BPW, repeated eight times to achieve dilutions from 10^{-5} to 10^{-8}. These dilutions were then spread on Chromogenic ECC Agar Plates [Edwards, Australia] and CHROMagarT ESBL Plates [Edwards, Australia] and incubated at 37 °C in an O_2 incubator for 16 to 18 h. Target colonies were then subcultured onto NA plates, incubated and transferred to MALDI plates as described for non-enriched cultures.

Isolates confirmed as *E. coli* then underwent DNA extraction using MagMAX CORE nucleic acid purification kits [ThermoFisher Scientific], following manufacturer's instructions. The culturing method used for each sequenced isolate can be viewed in Supplementary Data 2.

2.3. WGS, assembly and annotation

WGS was performed as previously described (Jarocki et al., 2021). Genomic DNA was prepared using Hackflex (Gaio et al., 2022) and sequenced on the Illumina NovaSeq platform. Draft genomes were assembled using Shovill v1.1.0 (https://github.com/tseemann/shovill) with the -trim option and default SPAdes assembler. Assembly statistics were obtained using assembly-stats v1.0.1 (https://github.com/sanger-pathogens/assembly-stats), and genomes were QCed using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc) and a Panaroo (Tonkin-Hill et al., 2020) packaged Mash wrapper script: any contaminated genomes or genome outliers with comparatively unusual numbers of genes or contigs were removed from all downstream analyses. The Prokka pipeline (Seemann, 2014) was used to generate automated annotations using the following options: -usegenus -genus Escherichia -compliant -mincontiglen 200 -addgenes. MOB-suite (Robertson and Nash, 2018) was used to assemble putative plasmid sequences using the mob_recon function, and Mlplasmids (Arredondo-Alonso et al., 2018) was used to corroborate plasmid- and chromosome-derived contigs.

2.4. Phylogenetic analysis

A Maximum-Likelihood phylogenetic tree was built using a core genome alignment (3,248,717 bp) via IQ-TREE 2 (Minh et al., 2020) v2.2.0. ModelFinderPlus (−m MFP) was used to determine the best-fit model (UNREST+FO + R4) and single branch testing was performed using ultrafast bootstrap (—B 1000). An additional *E. coli* ST131 SNP-based phylogeny was generated using 86 % (4,514,047 bp) of the *E. coli* ST131 strain EC958 genome (NZ_HG941718.1) as the reference, and the snplord snakemake pipeline (https://github.com/maxlcummins/pipelord/tree/master/snplord), which utilises snippy v4.3.6 (https://github.com/tseemann/snippy; default settings), Gubbins (Croucher et al., 2015) v2.3.4 (default settings), SNP-sites (Page et al., 2016) v2.4.1 (default settings), and FastTree (Price et al., 2010) v2.1.8 using -gtr and -nt options. Pairwise SNPs between isolates were calculated using snp-dists (https://github.com/tseemann/snp-dists) v0.8.2 and all trees were visualised using iTOL v6.5 (Letunic and Bork, 2021).

2.5. Genotyping

Sequence types (STs), serotypes and phylogroups were determined in silico using Achtman 7 multilocus sequencing typing (MLST) v2.0 (Larsen et al., 2012), SerotypeFinder v2.0 (Joensen et al., 2015), and Clermont Typing (Clermont et al., 2013), respectively. Plasmid MLST (pMLST), typing of *fimH* alleles, and ST131 virotyping were performed using pMLST v2.0 (Carattoli and Hasman, 2020), FimTyper v1.0 (Roer et al., 2017) and criteria established by Blanco et al. (2013), respectively. ABRicate v1.0.1 (https://github.com/tseemann/abricate) was used to screen for virulence-associated genes (VAGs), antimicrobial resistance genes (ARGs), chromosomal point mutations conferring AMR, and plasmid replicons using the following reference databases: Virulence Factor DataBase (VFDB) (Liu et al., 2019), the Comprehensive Antibiotic Resistance Database (CARD) (Alcock et al., 2020), ResFinder v4.4 (Florensa et al., 2022), PointFinder v4.1 (Zankari et al., 2017) and PlasmidFinder v2.1 (Carattoli and Hasman, 2020). Additional custom databases with AMR-associated insertion sequences (IS), class 1, 2 and 3 integron integrases, and additional ExPEC-associated VAGs were also utilised and can be accessed at https://github.com/maxlcummins/custom_DBs/blob/master/EC_custom.fa. ABRicate was executed using default parameters (minimal coverage ≥80 % and minimal identity ≥80 %), except when screening for truncated *intI1* genes where minimal coverage was lowered to >40 %. Identified ARGs were assigned functional groups using abritAMR (Sherry et al., 2022). AbritAMR was also used to identify metal resistance genes (MRGs) and Locus of Heat Resistance (LoHR) genes, the latter of which were then checked manually to determine genetic context. Prophage sequences were identified using PHASTER (Arndt et al., 2016), compared using Mauve (Darling et al., 2004), and visualised using EasyFig (Sullivan et al., 2011).

2.6. Statistical analyses

Classical (metric) multidimensional scaling (MDS) plots for VAGs and ARGs were created using R Studio v4.0.5 and the gglot2 v3.3.0 package using standard R package functions dist and cmdscale in conjunction with gene presence/absence matrices (1 = present; 0 = absent). Statistical tests were performed using pairwise Wilcoxon test with Benjamini–Hochberg *P*-value using the pairwise_wilcox_test from

the rstatix package.

3. Results

3.1. Phylogenetic analysis

A Maximum-Likelihood core genome phylogeny was built to ascertain the diversity of the 281 *E. coli* isolates derived from two WWTPs located in NSW, Australia (Fig. 1A), no significant differences were found between isolates originating from either WWTPs in terms of diversity or genotypes. A pangenome analysis found that the collection had a core genome consisting of 3326 genes (core = 3124; soft core = 202) and an accessory genome of 15,975 genes (shell = 2220; cloud = 13,755). The total collection consisted of 80 (28 %) ESBL-selected isolates (growth on selective media) and 201 (72 %) isolates grown on nutrient agar without antibiotics. From a whole genome sequencing perspective, little is known about the diversity of *E. coli* that are culturable from wastewater. By comparing the two cohorts in terms of ST distribution, it was observed that selecting for ESBL resistance significantly favours the selection of ST131 (z score = 8.5; *p*-value = 1.11e-17) which comprised 40 % ($n = 32$) of the ESBL-selected cohort versus 2 % ($n = 4$) of the non-selected group (Fig. 1B). Only six STs were prevalent (i.e., STs represented by three or more isolates) in the ESBL-selected cohort: ST131 (phylogroup B2; n = 32), ST410 (phylogroup C; $n = 8$; Clonal complex [CC] 23), ST3268 (phylogroup D; $n = 6$; CC38), ST69 (phylogroup D; $n = 5$), ST636 (phylogroup B2; n = 4) and ST1193 (phylogroup B2; n = 3; CC14) (Fig. 1B and C). Unsurprisingly, the non-antibiotic selected cohort had greater ST diversity and the top STs differed to the ESBL-selected group (top 3: ST58 (CC155) > ST10 (CC10) > ST95) (Fig. 1B); however, it is notable that several pandemic ExPEC linages were missing from the ESBL-selected group only, including ST12, ST58, ST73, ST95, ST117 and ST127.

3.1.1. SNP analysis of E. coli ST131 from wastewater

ST131 isolates dominated *E. coli* retrieved from ESBL-selective media ($n = 36$; 13 %). To determine the phylogenetic distribution of major ST131 clades, a Maximum-Likelihood phylogenetic tree was built using the complete genome of *E. coli* ST131 strain EC958 (NZ_HG941718) as a reference, the 36 ST131 isolates from this collection, and an additional 42 ST131 isolates (all available ST131 originating from Australia during the same time period [2018–2019] sourced from Enterobase [human = 38; avian = 4; metadata in Supplementary data 1]). The tree was constructed using 1952 SNPs from an 86 % alignment to the reference genome (4,514,047 Mbp) (Fig. 2). ST131 consists of four major clades – clade A (environment-associated), B (zoonosis-associated), C1 (human-associated) and C2 (human-associated carrying $bla_{\text{CTX-M-15}}$) (Li et al., 2021). In our wastewater collection, Clade A isolates predominated ($n = 27$; 75 %), followed by four isolates each in clades C1 and C2 and only one clade B isolate. Only three clade A wastewater isolates did not harbor $bla_{\text{CTX-M}}$ genes. The most common was $bla_{\text{CTX-M-27}}$ ($n = 15$), followed by $bla_{\text{CTX-M-15}}$ ($n = 9$) and one $bla_{\text{CTX-M-14}}$. The most common IncF RST in clade A wastewater isolates was F29:A-:B10 (n = 11; 31 %), the RST of pUTI89 and a major ExPEC F virulence plasmid RST. F29:A-:B10 is typically associated with clade A and in specific sublineages of clade B. Notably, 22 % (n = 8) of clade A isolates carried F1:A2:B20 plasmids, which are typically associated with clade C1. Of the 4 C1 isolates in our wastewater collection, 3 carried an F1:A2:B20 plasmid. (Fig. 2). In a broader phylogenomic study of Australian ST131 isolates, F1:A2:B20 was predominantly found in clade C1 isolates. A subset of clade A wastewater isolates also carried F1:A1:B16 plasmids ($n = 5$), and this RST was also present in a single clade C2 isolate in our wastewater collection. A SNP analysis identified several closely related clade A wastewater isolates to clinical isolates (pairwise SNP analyses in Supplementary Data 1), including a cluster of seven wastewater isolates (all virotype E) which differed by an average of 18 SNPs (range 15–23 SNPs) to human sepsis isolate SRR5936479 (BioSample: SAMN07510239). Similarly, a cluster of eight clade A wastewater isolates (all virotype A and carrying F1:A2:B20 plasmids) differed by an average of 27 SNPs (range 25–36 SNPs) to clinical isolate ESC_YB1643AA_AS (BioSample: SAMN20033273) and an average of 22 SNPs (range 20–31 SNPs) to pigeon (*Columba livia*) isolate SRR11341591 (BioSample: SAMN14395275). Curiously, one wastewater clade A isolate, DB_4_A4, was phylogenetically distinct to all other clade A isolates (Fig. 2) and differed by an average of 160 SNPs (the average SNP differences between other clade A isolates was 69 SNPs). While DB_4_A4 carried *fimH*41, typical of clade A ST131, it was O25b:H4, a serotype most frequently associated with clade C ST131. Isolate DB_4_A4 also carried dual *parC*-1aAB *gyrA*-1AB mutations, which up until recently (Li et al., 2021) were only reported in clade C.

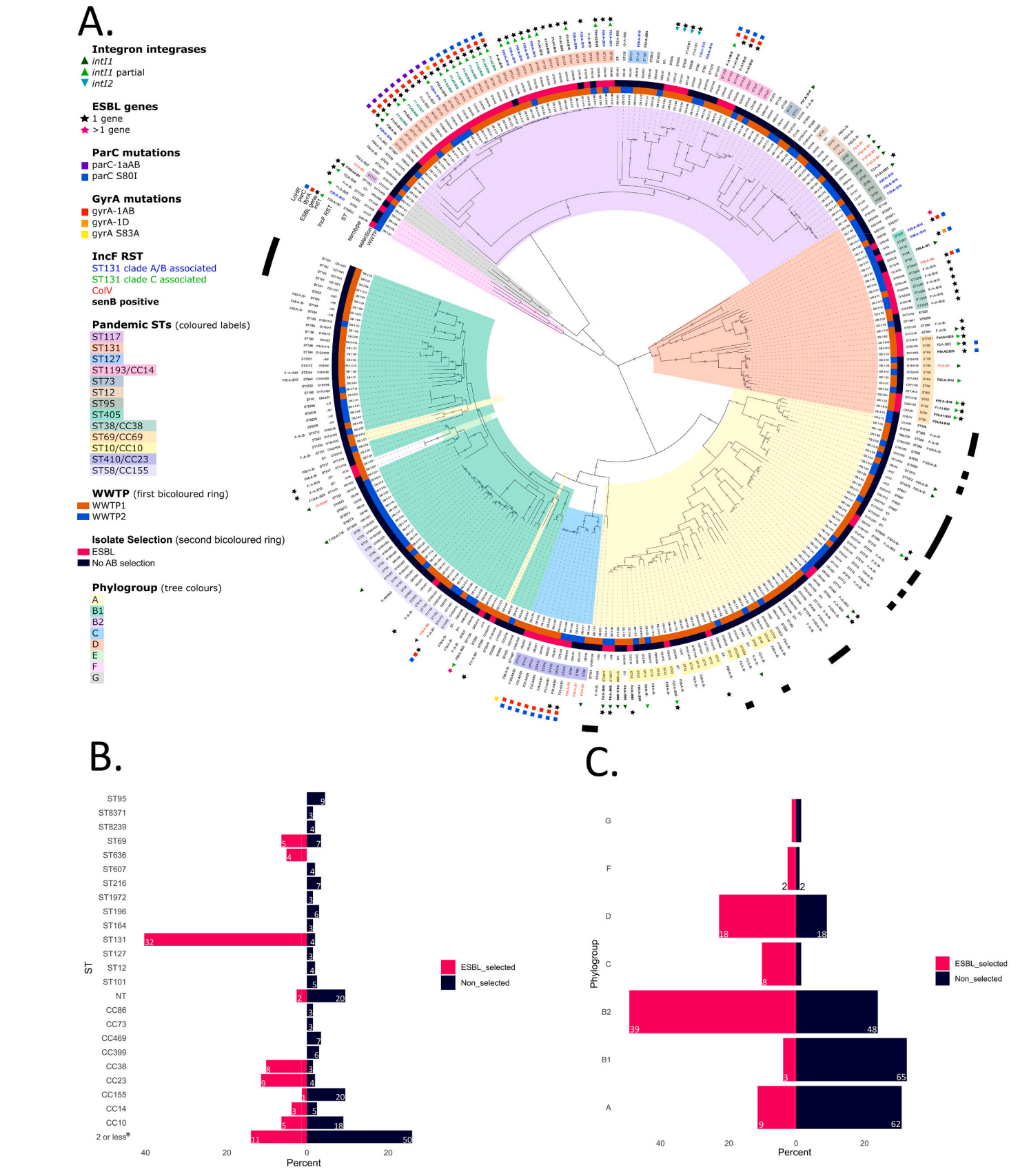
E. coli ST131 clade C isolates tend to be more clonal than clade A and clade B. The 16 C1 isolates (all virotype C) analysed here differed by an average of 42 SNPs (range 9–86 SNPs) and the closest wastewater to clinical isolate was DB_3_D12 to ESC_YB1712AA_AS (biosample: SAMN20033336), differing by 19 SNPs. Similarly, the 22 C2 isolates differed by an average of 64 SNPs (range 0–102 SNPs) and the closest wastewater to clinical isolate was D _5_C5 to SRR10126846, differing by 17 SNPs. Notably, two C2 wastewater isolates, DB_4_A2 and DB_3_H11, were serotype O12:H4, rather than the typical O25b:H4.

A genome-wide association study (GWAS) was performed to determine whether the ST131 wastewater isolates here possessed unique quantitative traits compared to clinical ST131 strains. We compared our 36 ST131 wastewater isolates to all ST131 human-sourced isolates available from Enterobase that originated from Australia and were collected during the same time frame as our collection (2019–2020) ($n = 34$; isolates with "ESC" prefix in Fig. 2). We found that 196 genes were statistically different between the two populations (scoary output in Supplementary Data 1); however, no gene had a > 85 sensitivity and specificity value. The highest sensitivity/specificity values were 80.5/85.3 for a group of six hypothetical proteins, which were significantly more present in wastewater isolates ($p = 8.42E-06$). Upon further scrutiny, the genes encoding these six hypothetical proteins were clustered together forming a 5810 bp region and additional Blast searches uncovered that all the genes encode for phage tail proteins.

3.2. Antimicrobial resistance profile

A total of 50 ARGs associated with HGT were identified (Fig. 3A; list of all identified ARGs [$n = 100$], including intrinsic efflux pumps, in Supplementary Data 2). We compared the ARG profiles of ESBL-selected isolates and non-selected isolates (Fig. 3B and D). The most common ARGs in the ESBL-selected cohort were *mph(A)* (macrolide resistance; $n = 45$; 56 %), $bla_{\text{CTX-M-15}}$ (ESBL; $n = 43$; 54 %) and *tetA* (tetracycline resistance; $n = 41$, 51 %), while the most common ARGs in the non-selected cohort were $bla_{\text{TEM-1b}}$ (narrow-spectrum β-lactamase; $n = 22$; 11 %), *tetA* ($n = 16$; 8 %) and *aph(3″)-Ib* (streptomycin resistance; $n = 15$; 7 %). While 91 % of ESBL-selected isolates carried $bla_{\text{CTX-M}}$ genes (43 $bla_{\text{CTX-M-15}}$, 26 $bla_{\text{CTX-M-27}}$, 2 $bla_{\text{CTX-M-14}}$, 1 $bla_{\text{CTX-M-3}}$, 1 $bla_{\text{CTX-M-55}}$) there were no $bla_{\text{CTX-M}}$ genes in the 201 non-selected isolates suggesting an overall low abundance of this gene family in the two targeted WWTPs. Selecting for ESBL resistance also selected for a higher prevalence of other ARGs, with isolates carrying an average of 6 ARGs (range 1–12) as opposed to 80 % ($n = 161$) of the non-selected isolates carrying no ARGs. Nevertheless, some critical important ARGs were identified in the non-selected group only, including *tet(X4)*, known to inactivate all tetracyclines including "last-resort" antibiotic tigecycline (Sun et al., 2019), found in one ST12264 isolate from phylogroup G (DB_2_H8) and Amp-C type ESBL gene $bla_{\text{CMY-2}}$ in two isolates (1 ST2190 and 1 novel ST) from phylogroup D.

Point mutations in genes conferring resistance to beta-lactams (*ampC*, *ftsI*, *marA*, *cyaA*), fosfomycin (*glpT*, *ptsI*, *uhpT*), and fluoroquinolone (*gyrA*, *parC*, *parE*) were also identified (Supplementary Data 2). Markedly, 87 % ($n = 244$) of all wastewater isolates possessed the



(caption on next page)

Fig. 1. Phylogeny and plasmid replicon carriage in *E. coli* from two Australian WWTPs. A) Maximum-Likelihood tree inferred from a core genome alignment. Bootstrap values >0.9 are shown as grey dots on branches. Isolates marked with * (n = 5) have been flagged as having potential mutations in genes determining phylogrouping, as they do not match their neighbouring mash group (mash group = B1, assigned phylogroups A (n = 3) and E (*n* = 2)). B) Mirror bar graph illustrating the distribution of STs across ESBL-selected isolates (pink) and non-antibiotic selected isolates (raisin) as a percentage of each respective cohort. The number of isolates per ST/CC is shown in the corner of each bar. CC86 includes ST453 (n = 2) and ST1490 (*n* = 1). CC73 includes ST73 (n = 2) and ST355 (n = 1). CC469 includes ST162 (n = 1), ST517 (n = 2) and ST5073 (n = 4). CC399 includes ST399 (n = 5) and ST635 (n = 1). CC38 includes ST38 (n = 3, 1 in ESBL-selected), ST3268 (n = 6, all ESBL-selected) and ST2003 (n = 1, ESBL-selected). CC23 includes ST88 (n = 3), ST295 (n = 2, 1 ESBL-selected) and ST410 (*n* = 8, all ESBL-selected). CC155 includes ST58 (*n* = 11), ST155 (n = 3), ST683 (n = 2), ST616 (n = 1, ESBL-selected), and ST1049, ST2170 and ST5487 (one each). CC14 includes ST537 (n = 4) and ST1193 (n = 4; 3 ESBL-selected). CC10 includes ST10 (*n* = 12, 2 ESBL-selected), ST43 (n = 2, 1 ESBL-selected), ST48 (n = 2), ST218 (n = 2), ST1286 (n = 1), ST3877 (n = 2, both ESBL-selected) and ST5334 (n = 2). C) Mirror bar graph illustrating distribution of phylogroups across ESBL-selected isolates (pink) and non-antibiotic selected isolates (raisin) as a percentage of each respective cohort. The number of isolates per phylogroup is shown in the corner of each bar.

fosfomycin resistance E448K mutation in *glpT*. The only isolates that did not possess this mutation were every CC10 isolate (*n* = 23), three novel ST isolates (DB_5_B8, DB_4_E2 and DB_2_A10) that were nestled between CC10 isolates in our phylogenetic analysis (Fig. 1A), all ST1193 isolates (n = 4, though these possessed the fosfomycin resistance E350Q mutation in *uhpT*), and most ST58 isolates (*n* = 7/11). Twenty-two isolates possessed a combination of *glpT* E448K and either *uhpT* E350Q (*n* = 20) or *ptsI* V25I (n = 2). Only ST131 isolates (*n* = 34; 94 % of ST131s) possessed all three fosfomycin resistance mutations. Fluroquinolone resistance associated mutations in *gyrA, parC* and/or *parE* where found in 27 % (*n* = 75) of isolates. In addition to the expected ST131 isolates, another 24 isolates had mutations conferring fluroquinolone resistance, including all ST1193 isolates (n = 4; *gyrA*-1AB, *parC* S80I, *parE* L416F) and all ST410 isolates (*n* = 8; *gyrA*-1AB, *parC* S80I, *parE* S458A).

Regardless of whether the isolates were selected for ESBL resistance or not, an MDS analysis demonstrated that the B2 phylogroup could be clearly delineated by ARG profile from other phylogroups, and that most ST131 isolates formed a separate cluster (Fig. 3C).

3.3. Virulence profile

A total of 283 virulence-associated genes (VAGs) were identified in this wastewater collection (selection in Fig. 3C; Supplementary Data 2). *E. coli* pathotypes are based on VAG carriage and we found that 95 (34 %) isolates satisfied molecular definitions of ExPEC (≥2 of 5 ExPEC markers, i.e., *afa/draBC*, *iutA*, *kpsMII*, *papAH* and/or *papC*, *sfa/focDE*, and/or ≥ 3 of 4 UPEC markers, i.e., *chuA*, *fyuA*, *vat*, *yfcV*) (Johnson et al., 2015; Johnson and Stell, 2000) (Fig. 4D; Supplementary Data 2). We found using both definitions, rather than one, captured most isolates belonging to pandemic ExPEC STs (ST10, ST117, ST1193, ST12, ST127, ST131, ST38, ST405, ST410, ST58, ST69, ST73, ST95) (*n* = 106 isolates of pandemic STs vs 95 molecularly defined ExPEC). The *Yersinia* High Pathogenicity Island (HPI), genotyped by the presence of *irp2* and *fyuA* (Tu et al., 2016), is correlated with blood stream infections in mouse models (Royer et al., 2023) and was found in 121 isolates, including 87 ExPEC. In addition to ExPEC, we identified eight atypical enteropathogenic *E. coli* (aEPEC; eae^{+}, bfp^{-}, stx^{-}), three enterotoxigenic *E. coli* (ETEC; heat-labile toxin^{+} (*eltAB*) and/or heat-stable toxin^{+} (*sta1*)) and one enteroaggregative *E. coli* (EAEC; aaf^{+}, $aggR^{+}$).

Notably, one ETEC isolate, DB_3_D11, was ST69, a well-known ExPEC sequence type. DB_3_D11 constitutes a hybrid pathotype, carrying ETEC heat-liable toxin encoding genes *eltA* and *eltB*, as well qualifying as ExPEC by molecular definition, hosting many ExPEC associated VAGs including *fyuA* (siderophore yersiniabactin receptor), *iutA* (siderophore aerobactin receptor), *kpsMII* (K1 capsule), and *papA* (uroepithelium-specific P fimbriae). Conversely, while two aEPEC isolates carrying Locus of Enterocyte Effacement (LEE) pathogenicity island marker gene *eae* (encoding intimin) were ST10, another well-known pandemic ExPEC sequence type, these two isolates did not carry ExPEC associated genes.

MDS analyses using VAG presence/absence found that isolates did not distinctly cluster by pathotypes though ExPEC isolates were predominately situated to the right and similarly some clustering by phylogroup was apparent (Fig.4A), with phylogroup A and B1 forming clusters to the left and B2 to the right. The phylogroup A isolates that deviated most belonged to CC10 (Fig. 4A, red points on the right of the dotted line). Overall, phylogroup A isolates carried significantly fewer VAGs (Fig. 4B; *p* values ranging from *p* = 1.93E-02 [A vs G], to *p* = 1.51E-21 [A vs B2]). Conversely phylogroup B2 isolates carried significantly more VAGs (A vs B2 p = 1.51E-21; B1 vs B2 *p* = 2.87E-22; C vs B2 *p* = 1.15E-05 and D vs B2 *p* = 5.13E-04). Due to high prevalence of ST131, ExPEC outnumbered *E. coli* without a defined pathotype in the ESBL-selected cohort, nevertheless ExPEC comprised 23 % of the non-selected cohort and was the only prominent pathotype in each collection.

3.4. Metal, biocide, and heat resistance

The transmissible Locus of Stress Tolerance (tLST) is a genomic region that confers an extreme heat-resistance phenotype along with greater resistance to chlorination and other advanced oxidants (Wang et al., 2020; Yu et al., 2022). Here we identified 40 isolates harbouring a tLST (Fig. 1A), all of which were subtype 1 (tLST1) (Zhang and Yang, 2022) (Fig. 5C). These isolates were from phylogroup A (*n* = 34) and B1 (*n* = 6) and belonged to 15 STs, including three CC10 (1 ST10, 1 ST48, 1 ST218), five ST101, six ST216, one ST226, six CC399 (5 ST399, 1 ST635), two ST401, one ST409, four ST607, two ST1316, three ST1972, one ST4088 and 1 ST5295 isolate, plus 5 isolates with novel STs (3 sharing the same serotype O173:H27). The tLST isolates were generally unremarkable in terms of VAG carriage, except that two ST399 isolates (one O86:H30, one O1:H30) and one ST226 (O85:H10) isolate carried HPI markers (*fyuA* and *irp2*). Similarly, isolates with tLST carried next to no acquired ARGs with only a singular isolate carrying one ARG (*qnrB5* in an ST101 isolate), though multi-drug efflux pumps, such as *cpxA, CRP, mdfA, marA,* etc. were present in all tLST isolates and ubiquitous in all 281 isolates (Supplementary Data 2). The locus was not identified in any ESBL-selected isolate. However, tLST isolates were found to carry significantly more metal resistance genes (Fig. 5B), conferring resistance to arsenic, copper, mercury, nickel, and silver (Fig. 5A).

3.5. MGEs

3.5.1. General

In *E. coli,* plasmids influence evolutionary trajectories, the development of pathotypes, and AMR (Cummins et al., 2022; Johnson et al., 2016; Johnson and Nolan, 2009; Reid et al., 2022). Here we looked at the distribution of plasmid replicons (Fig. 6A; Supplementary Data 2) and F plasmid replicon sequence types (RST) (Fig. 6B; Supplementary Data 2) between isolates selected for ESBL-resistance versus a non-selected cohort devoid of ESBL genes. In both collections, IncF plasmids were predominant, however their RSTs differed. Out of the 69 IncF RSTs identified, only 6 were shared – F29:A-:B10, F-:A1:B10, F1:A-:B23, F4, A-:B65, F30:A-:B- and those non typable (F-:A-:B-) (Fig. 6B).

3.5.2. Putative AMR plasmids and the genetic context of $bla_{CTX\text{-}M}$ genes

Using MOB-suite, 1031 draft plasmid sequences were assembled, of

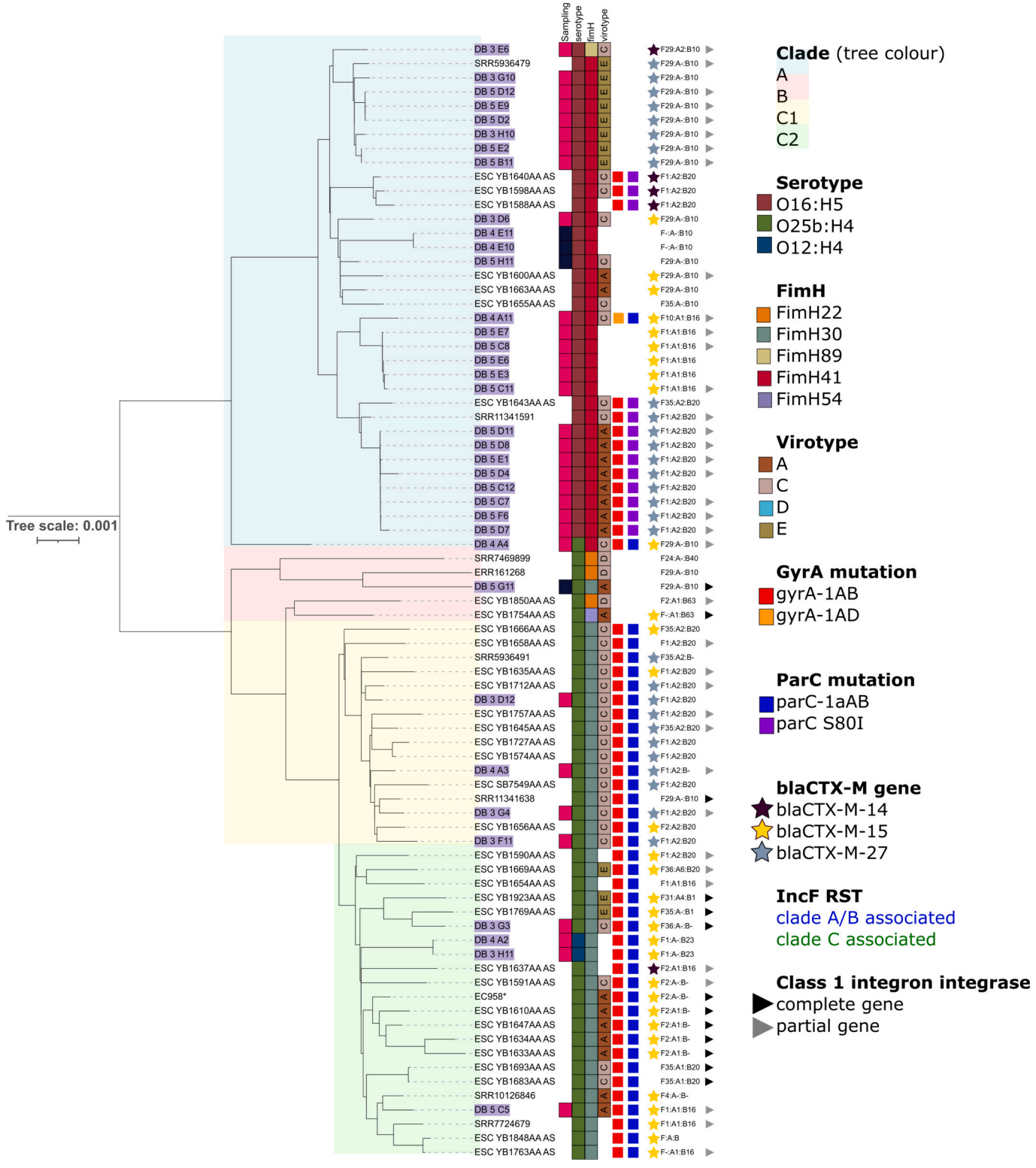


Fig. 2. Phylogeny of *E. coli* ST131 from wastewater. Maximum-Likelihood SNP-based tree using ST131 reference strain EC958 as reference (marked with *). Bootstrap values >0.95 are shown as black dots on branches. Isolates from this collection are marked by purple tip labels, human-sourced isolates have unmarked tip labels and avian isolate tip labels are coloured tan.

which 137 (13 %) carried ARGs. These putative AMR plasmids predominantly carried IncF replicons ($n = 87$; 64 %), but ARGs were also located on IncI1 ($n = 4$), IncK2/Z ($n = 9$), Col ($n = 34$, 30 of which were col(156)), IncX1 ($n = 1$), IncX4 (n = 1) and IncY (n = 1) putative plasmids. The singular *tet(X4)* gene was situated on a IncF plasmid (F2: A-:B-) along with *floR* (chloramphenicol), *lnuF* (lincomycin) and *aadA2* (streptomycin). Notably, 62 (45 %) of putative AMR plasmids also carried VAGs, predominately either *senB* ($n = 39$) or aerobactin ($n = 20$; *iucABCD*, *iutA*) (Supplementary Data 2). Similarly, the putative plasmids that housed $bla_{\text{CTX-M}}$ genes (n = 34) were mainly IncF and frequently

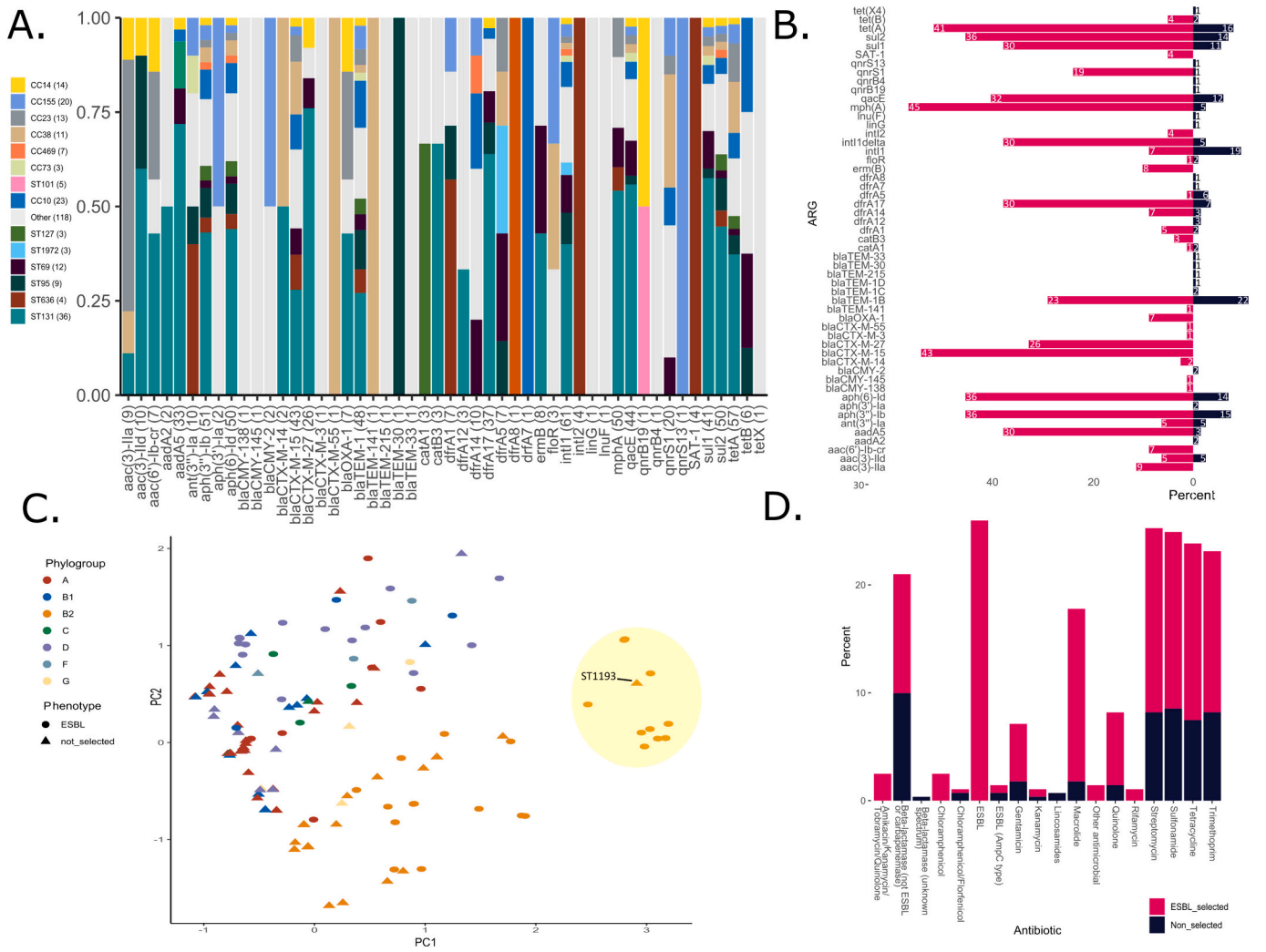


Fig. 3. AMR profile of 281 wastewater *E. coli* isolates. A) Percent stacked bar chart illustrating the distribution of ARGs associated with HGT (gene counts in brackets) among isolates by ST or CC (number of isolates per ST/CC in brackets). B) Mirror bar plot illustrating the distribution of ARGs in ESBL-selected (pink) versus non-selected (raisin) cohorts as a percentage of each collection. The gene counts are shown at the end of each bar. C) MDS chart illustrating clustering by ARG presence/absence using all identified ARGs (n = 89). Coloured by phylogroup, shapes represent the ESBL and non-selected cohorts. A ST131 cluster (n = 21) is highlighted in yellow. D) Bar chart illustrating the percent of isolates carrying ARGs that confer resistance to specific antibiotics.

carried additional ARGs and VAGs (Fig. 7A). While $bla_{CTX\text{-}M}$ genes are most frequently reported as being plasmid-mediated, we found 33 situated chromosomally in the following STs: ST10 [n = 1 $bla_{CTX\text{-}M\text{-}15}$], ST1193 [n = 1 $bla_{CTX\text{-}M\text{-}15}$], ST131 [n = 9 $bla_{CTX\text{-}M\text{-}15}$, n = 1 $bla_{CTX\text{-}M\text{-}14}$, n = 7 $bla_{CTX\text{-}M\text{-}27}$], ST1722 [n = 1 $bla_{CTX\text{-}M\text{-}15}$], ST3268 [$n$ = 6 $bla_{CTX\text{-}M\text{-}15}$], ST38 [$n$ = 1 $bla_{CTX\text{-}M\text{-}14}$], ST43 [n = 1 $bla_{CTX\text{-}M\text{-}27}$], ST616 [n = 1 $bla_{CTX\text{-}M\text{-}15}$], and ST636 [$n$ = 3 $bla_{CTX\text{-}M\text{-}15}$]) and while they occurred in different genetic context, all chromosomal $bla_{CTX\text{-}M}$ alleles were found associated with IS*Ecp1* (examples provided in Fig. 5B), except for when a contig break occurred directly upstream to $bla_{CTX\text{-}M}$. Notably, in all clade A ST131 isolates with chromosomal $bla_{CTX\text{-}M\text{-}27}$ (n = 7), the gene was integrated into *gspD,* a characteristic of a global clade A sublineage with stable $bla_{CTX\text{-}M\text{-}27}$ integration (Biggel et al., 2023). Chromosomal $bla_{CTX\text{-}M}$ genes have recently been correlated with the presence of F virulence plasmid, pUTI89, at least in the Australian *E. coli* ST131 population (Li et al., 2021), and indeed we found that 61 % (n = 20) of isolates with identified chromosomal $bla_{CTX\text{-}M}$ genes carried pUTI89-like plasmids (i.e., F29:A-:B10).

Also of note, in one clade A ST131 isolate (DB_3_D6) we found that $bla_{CTX\text{-}M\text{-}15}$ was carried within an intact phage element situated 6785 bp from a RepFIB plasmid replication encoding gene (Fig. 6C). DB_3_D6 carries F29:A-:B10, the RST of pUTI89, however it also carries an additional IncF replicon – IncFIB (H89-PhagePlasmid) – which, as the name implies, is found on phage-like plasmids. An NCBI BLAST search found that the DB_3_D6 intact prophage sequence had a strong match (≥84 % coverage, >98 % identity) to four *E. coli* phage-like plasmids. While all four compared phage-like plasmids carried $bla_{CTX\text{-}M\text{-}15}$, only in two, KY515224.1 (wildlife; USA) and LC056430.1 (water; India) was the gene situated within regions homologous to the DB_3_D6 prophage sequence (Fig. 7C).

3.5.3. Class 1 integrons

Class 1 integrons play a key role in AMR due to their ability to capture and express diverse ARGs. While not themselves mobile, class 1 integrons are frequently mobilised by adjacent transposons and their presence is considered a proxy for an MDR profile and for anthropogenic pollution (Gillings et al., 2015). We identified 26 full length *intI1* and 35 with IS*26*-induced truncations (total 22 %). Truncated *intI1* genes were more common in ESBL-selected isolates (30 vs 5) while full length versions were present in equal proportions (i.e., 9 % of isolates in both collections). We observed 10 primary structures of which the most common was *intI1-dfrA17-aadA5-qacEΔ1-sul1*-ORF-*chrA-padR*-IS*6100-mphR-mrx-mphA* (n = 29, Fig. 7D), conferring resistance to trimethoprim, streptomycin, quaternary ammonium compounds, sulphonamide,

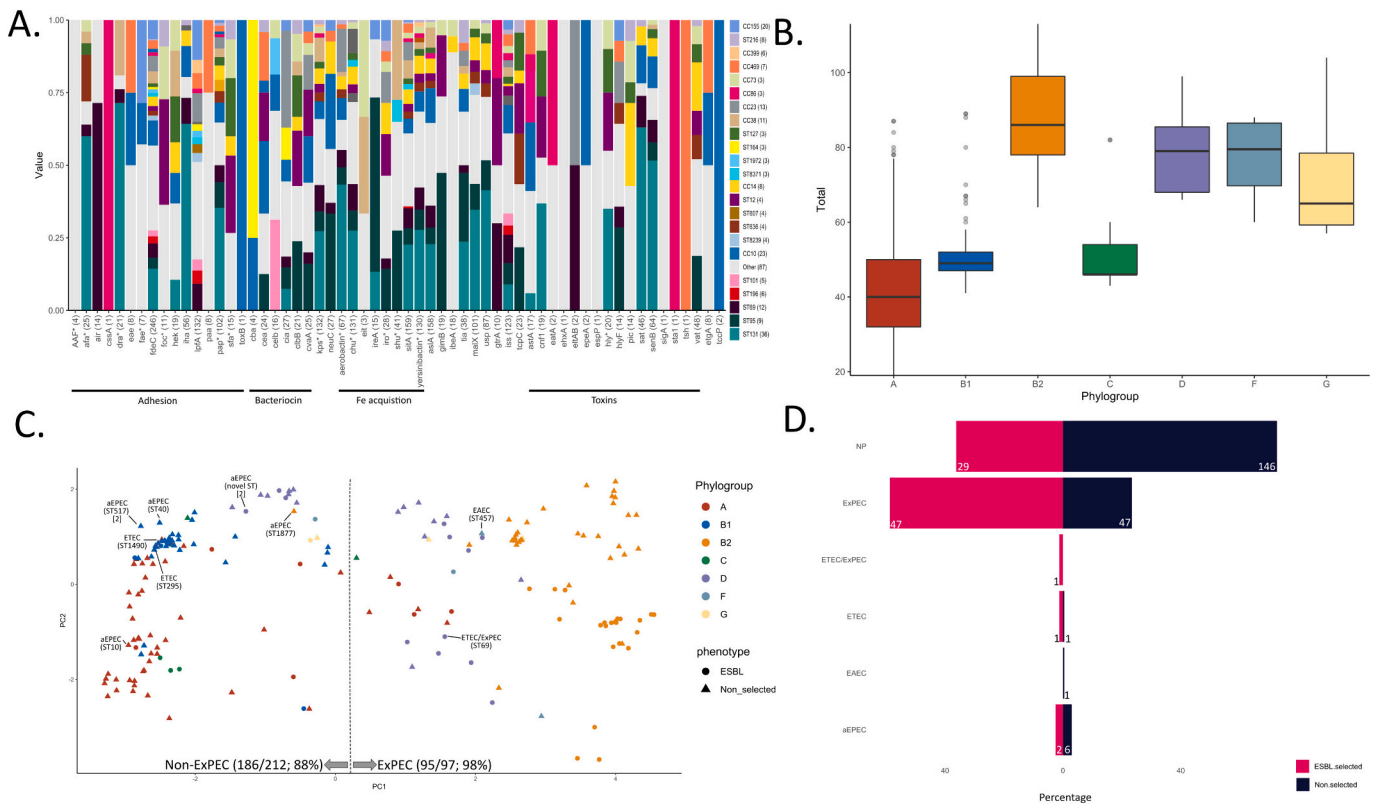


Fig. 4. Virulence profile of 281 wastewater *E. coli* isolates. A) MDS plot illustrating clustering by VAG presence/absence. Coloured by phylogroup, shapes represent WWTP origin. Pathotypes other than ExPEC indicated by text. B) Boxplot illustrating the distribution of VAG counts across phylogroups. C) Percent stacked bar chart illustrating the distribution of VAGs (gene occurrence in brackets) among isolates by ST or CC (number of isolates per ST/CC in brackets). D) Mirror bar chart illustrating the percent of isolates according to pathotypes. NP = no pathotype.

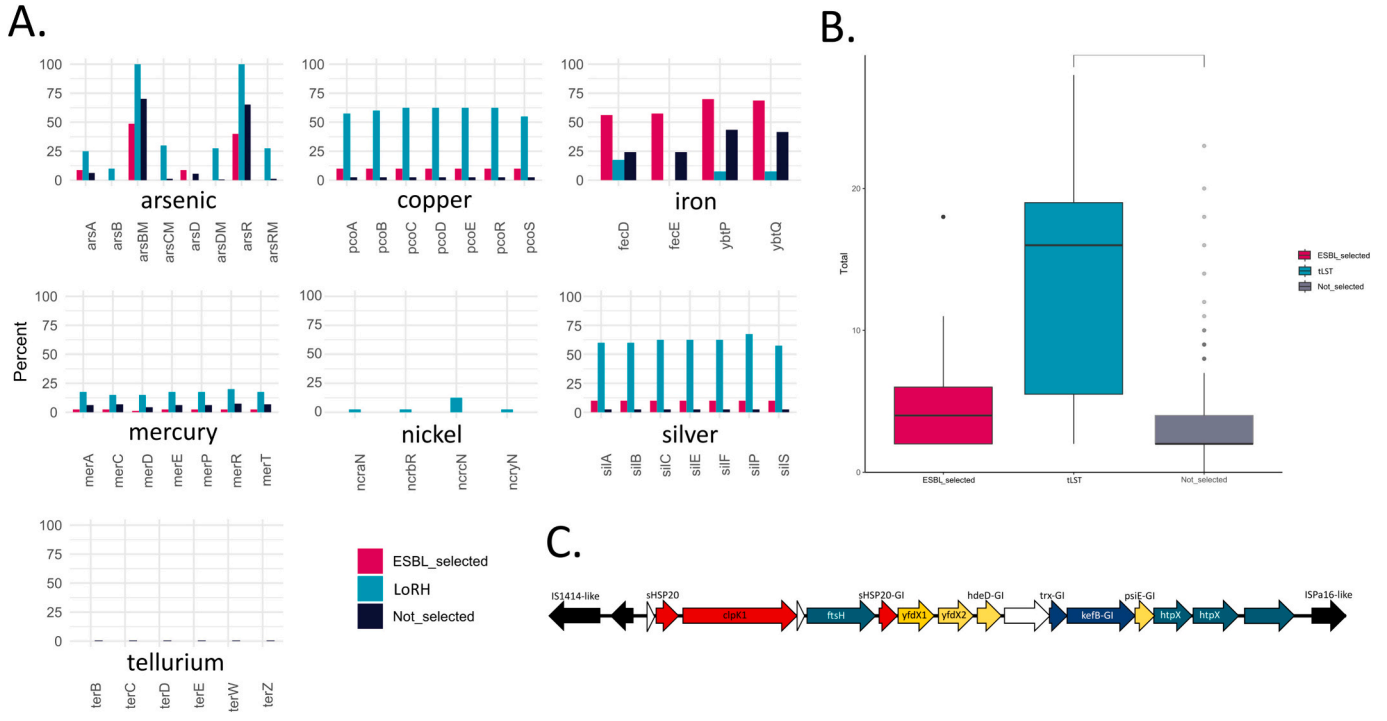


Fig. 5. Isolates with LoHR carry more metal resistance genes. A) Bar graphs illustrating the percent of isolates carrying metal resistance genes (percent of each respective cohort – isolates with LoHR, ESBL-selected isolates, Non-selected isolates). B) Box plot illustrating the number of metal resistance genes across the three cohorts. C) Schematic of the tLST. The loci was found to be flanked by IS elements with limited homology to those deposited in ISFinder database – Left: ISPa16: 70.81 % (*Pseudomonas* origin), ISPpu19: 48.61 % (*Pseudomonas* origin); Right: IS1414: 79.15 % (*E. coli* origin)

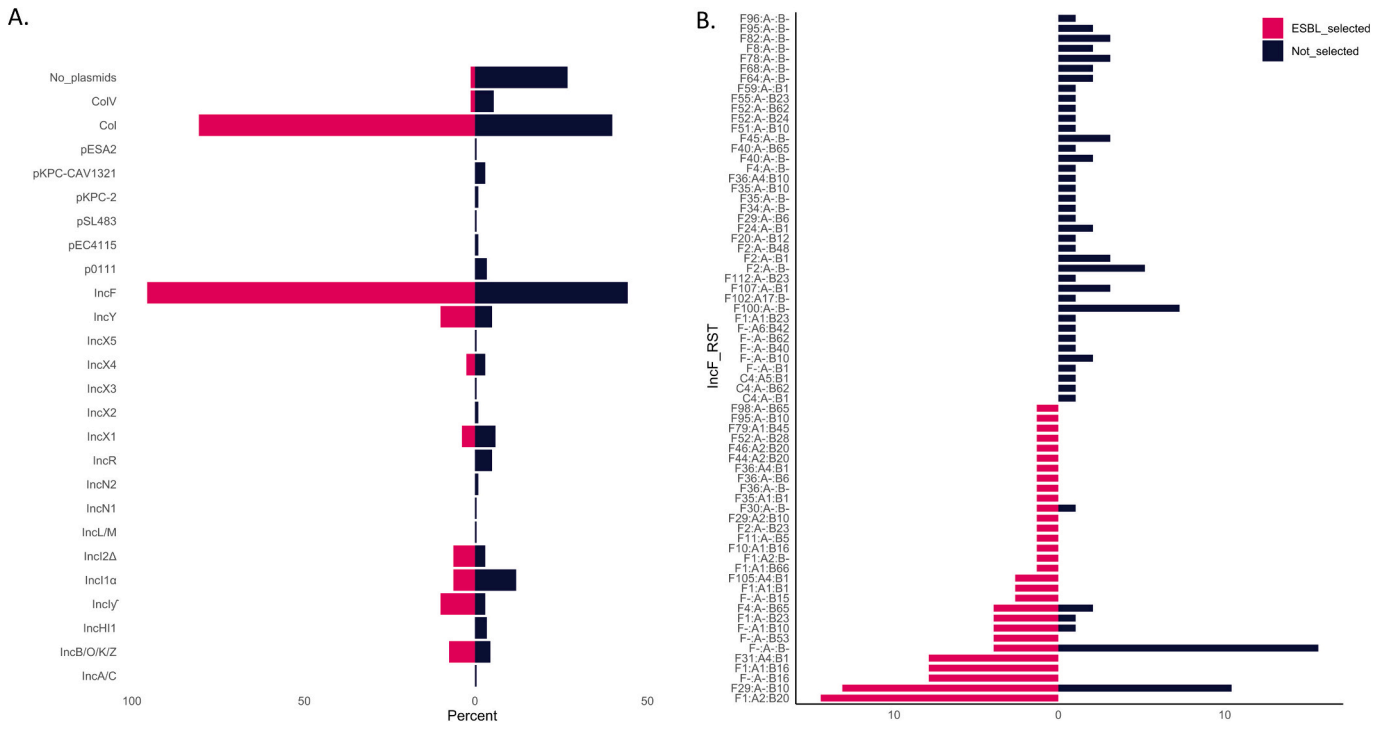


Fig. 6. Distribution of plasmid replicons across ESBL-selected and non-selected isolates. A) Mirror bar plot illustrating the distribution of plasmid replicons in ESBL-selected (pink) versus non-selected (raisin) cohorts as a percentage of each collection. B) Mirror bar plot illustrating the distribution of IncF plasmid RSTs in ESBL-selected (pink) versus non-selected (raisin) cohorts as a percentage of each collection.

chromate, and macrolide. This complex resistance region was prevalent in ST131 isolates (n = 20) but also present in ST1164, ST69, ST1136, ST10, ST6870, ST43, ST1193 and ST7401 isolates. Given that *intI1* truncations can serve as potential epidemiological markers (Li et al., 2021; Wyrsch et al., 2022) it was worth noting that, in the most common structure, full length *intI1* was only found in one ST131 and one ST1164. The remaining 27 instances had either $intI1\Delta_{762}$ (*n* = 12; ST131, ST69, ST7401, ST1193), $intI1\Delta_{600}$ (n = 7; ST131), $intI1\Delta_{699}$ (n = 1; ST43), $intI1\Delta_{648}$ (n = 2; ST69, ST1136), and $intI1\Delta_{564}$ (n = 3; ST131). We also identified the typical class 2 integron structure, *intI2-dfrA1* (trimethoprim resistance)-*SAT-1* (streptothricin resistance)-*ant(3″)-IIa* (streptomycin resistance), in four ST636 isolates (all ESBL-selected).

3.5.4. Putative tLST-positive plasmids

The tLST was identified within putative plasmid sequences in 14 isolates (35 % of $tLST^{+}$ isolates). Eight isolates contained plasmid sequences that matched closest to plasmid CP026404 from *E. coli* strain ECONIH4 (AA057; average mash distance 0.043) and two isolates matched closest to plasmid CP008907 from *Enterobacter cloacae* strain ECR091 (AA425; average mash distance 0.049). Both of these plasmids were IncF (CP026404 = F100:A-:B-; CP008907 = F-:A-:B-), however plasmid sequence alignments demonstrated that, apart from the tLST, the putative plasmid sequences shared little homology to these reference plasmids (Fig. 8). For the remaining four isolates, the closest matched reference plasmids either did not contain a tLST, or the putative plasmid sequences were assigned as novel (i.e., not seen previously in the MOB-suite database). Future long-read sequencing is required to validate these initial observations.

4. Discussion

E. coli is the causative agent of a spectrum of infections, ranging from urinary tract infections to severe cases of gastroenteritis and systemic diseases. *E. coli* is also one of the primary ESBL-producing organisms within *Enterobacterales* and the leading pathogen for deaths associated with AMR (Murray et al., 2022; Pitout and Laupland, 2008). In Australia, resistance in *E. coli* to the extended-spectrum beta lactam cefotaxime has been steadily increasing and, as of 2021, cefotaxime resistance ranks close to the median rate for European countries (12.5 %) (AURA, 2023). While there is undoubtedly merit in the continued surveillance of ESBL-producing *E. coli*, it is important to investigate the biases such selection-criteria can generate. By performing a comparative genomic analysis of 80 ESBL-selected *E. coli* isolates and 201 non-selected *E. coli* isolates from influent wastewater in NSW, Australia, we found major variances between the two cohorts in terms of diversity, AMR profiles, and the carriage of VAGs, plasmids, and the tLST.

4.1. Diversity is lower in ESBL-selected isolates and heavily skewed towards ST131

E. coli exhibits remarkable genomic diversity, attributable to its extensive genome plasticity, global distribution, and presence in a variety of ecological niches. While its core genome encompasses approximately 2000 genes, pangenome analyses have uncovered an accessory genome comprising tens of thousands of genes, with an estimated addition of 26 new genes for each genome sequenced (Denamur et al., 2021; Kallonen et al., 2017; Touchon et al., 2020). This expansive genetic repertoire, combined with a high frequency of horizontal gene transfer, positions *E. coli* as a critical reservoir for the dissemination of various genes, including ARGs and VAGs. Consequently, sampling protocols that potentially reduce genetic diversity can impede efforts to monitor and address emerging resistance or virulence traits in *E. coli* populations. In our study, we observed a substantial reduction in diversity among ESBL-selected isolates compared to their non-selected counterparts, with 3.65 times fewer STs identified (23 STs + 2 non-typable isolates vs. 84 STs + 20 non-typable isolates). While the difference in diversity may in part be due to a smaller number of isolates in our ESBL-selected cohort, another study using equal numbers of non-ESBL and ESBL selected *E. coli* also reported decreased diversity in the latter (116 STs vs 64 STs) (Raven et al., 2019). The diversity of ESBL-

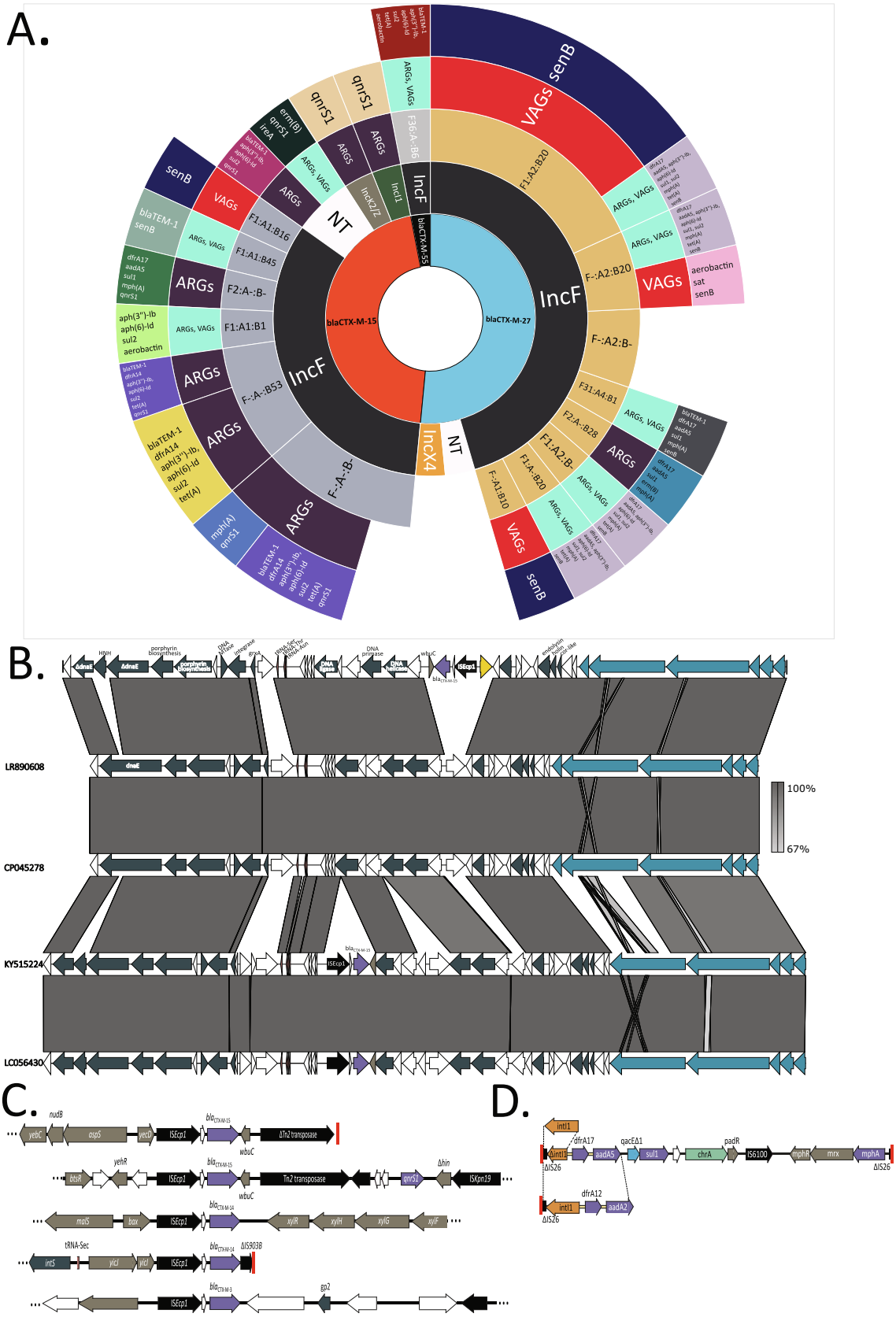


Fig. 7. Genetic context of ARGs. A) Sunburst graph representing the IncF RSTs of predicted plasmids carrying *bla*$_{CTX-M}$ genes and any additional ARGs and VAGs. B) The genetic context of *bla*$_{CTX-M-15}$ within ST131 isolate DB_3_D6 compared to four phage-like plasmids. C) Examples of *bla*$_{CTX-M}$ genes in various genetic contexts but always in association with IS*Ecp1*. D) Schematic of the most common Class 1 integron structure in this collection with adjacent macrolide resistance operon (*mphR-mrx-mphA*).

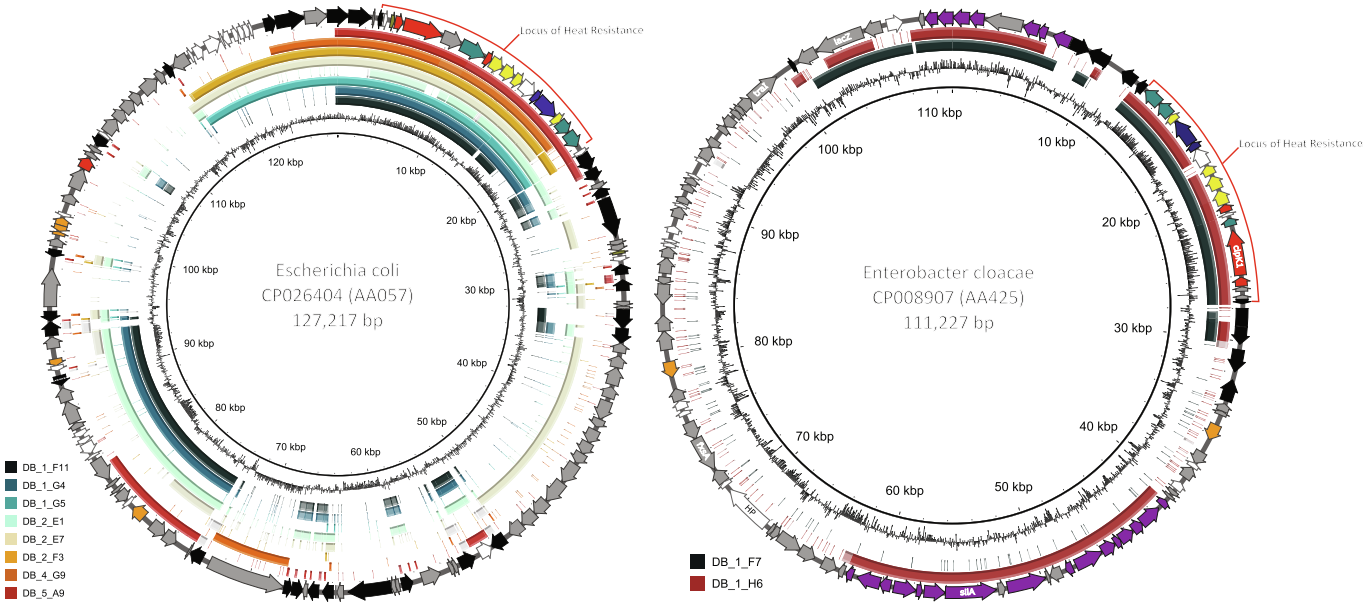


Fig. 8. Putative tLST plasmids. Left: Draft plasmid sequence alignments to reference *E. coli* IncF plasmid CP026404 containing the tLST. Right: Draft plasmid sequence alignments to reference plasmid CP008907 from *E. cloacae*.

producing *E. coli* is generally lower than that of non-ESBL-producing strains as antibiotic usage exerts selective pressures that promote the proliferation of specific clones, which are more likely to harbor ESBL genes like $bla_{\text{CTX-M}}$. This clonal expansion reduces genetic variability, as evidenced by the dominance of certain STs, such as ST131, among ESBL-producing *E. coli* (Day et al., 2016; van Hout et al., 2020). Indeed, we also observed that selection for ESBLs exhibited a significant selection bias for ST131 (*p*-value 1.11e-17), which constituted 40 % of the ESBL-selected cohort compared to 2 % in the non-selected group. ST131 has consistently emerged as the predominate ST detected in wastewater when ESBL-selection is applied, across diverse regions such as Czechia (Davidova-Gerzova et al., 2023; Dolejska et al., 2011), England (Raven et al., 2019), Japan (Sekizuka et al., 2022; Shibuki et al., 2023; Tanaka et al., 2019), Norway (Jørgensen et al., 2017), Portugal (Tavares et al., 2020) and now Australia. Contrastingly, studies not employing antibiotic selection report considerably lower ST131 prevalence in wastewater. In Canada, only 1.86 % of 3762 *E. coli* wastewater samples harboured ST131 (Finn et al., 2020), and in Norway, <1 % of 3123 wastewater *E. coli* isolates were ST131 (Paulshus et al., 2019). A study in regional NSW Australia, found that ST131 constitutes 2–5 % of *E. coli* in normal human faecal flora (Kudinha and Kong, 2022), aligning with the percentage observed in our non-selected municipal wastewater influent isolates. However, there is currently a lack of data on the clade structure of commensal ST131 (ST131 has four major clades: A, B, C1 and C2), which is important as clade A isolates have been predominantly identified in wastewater samples from our study (75 % of ST131s), as well as in Canada (Finn et al., 2020), China (Zhong et al., 2015), and Switzerland (Biggel et al., 2023). The predominance of this clade in wastewater raises questions about whether it represents ST131 in typical asymptomatic faecal flora or if it is particularly well adapted to survival in wastewater environments.

The prevalence of ST131 as the principal ESBL-producing clone in wastewater is well-documented globally. Consensus also emerges on other dominant STs in WGS studies focused on ESBL-selective samples, notably ST10, ST38, and frequently ST648 (Jørgensen et al., 2017; Raven et al., 2019; Sekizuka et al., 2022; Tanaka et al., 2019; Zahra et al., 2018). In our study, while ST10 and ST38 were present in the ESBL-selected cohort, the most prevalent STs, aside from ST131, were ST410 and ST3268, with a notable absence of ST648 in both groups. ST410 is phylogroup C (Phylogroup A in older Clermont typing scheme) and recognized as an emerging high-risk clone, often linked to resistance against key antimicrobials such as fluoroquinolones, 3rd generation cephalosporins, and carbapenems (Roer et al., 2018). Our ST410 isolates exhibited dual *gyrA/parC* mutations, indicative of fluoroquinolone resistance, but lacked carbapenemase genes (as did all our studied isolates). Intriguingly, while two ST410 isolates carried the ESBL gene $bla_{\text{CTX-M-15}}$, the ESBL resistance phenotype of the remaining six could not be genotypically correlated, suggesting the potential presence of uncharacterized ESBL genes. Regarding ST3268, this rare phylogroup D ST (only 58 isolates reported on Enterobase as of 20th January 2024) has nevertheless been identified in ESBL-selected samples from diverse locations, including Sweden (gulls) (Atterby et al., 2017) and India (wastewater) (Paul et al., 2020), and belongs to CC38 (like ST38), a globally distributed high-risk clonal complex often linked to UTIs and ESBL production. Our ST3268 isolates harboured $bla_{\text{CTX-M-15}}$, alongside *qnrS1* (fluoroquinolone) and *tetA* (tetracycline) genes.

For non-selected wastewater, while fewer studies exist, ST399 has emerged as a common ST in England (Raven et al., 2019) and Australia (septic tanks) (Behruznia and Gordon, 2022), and is considered a naturalised *E. coli* ST. While we identified five ST399s, the most common STs in our non-selected isolates belonged to CC155 ($n = 19$), particularly ST58 ($n = 11$), a globally disseminated ExPEC clone increasingly associated with sepsis (Reid et al., 2022). Notably, unlike most pandemic ExPECs which belong to pathogen-associated phylogroups B2 and D, ST58 belongs to the environmental/commensal-associated phylogroup B1. Incidentally, phylogroup B1 and A predominated in the non-selected cohort (32 % and 31 % respectively), which agrees with the characterisation of these phylogroups as generalists (Touchon et al., 2020). Conversely, the ESBL-selected group was primarily composed of ExPEC-associated phylogroups (Touchon et al., 2020), with B2 representing 49 % and D accounting for 23 % of isolates.

Alongside diminished diversity, a bias towards ST131, and the variation in top STs and phylogroups identified, the ESBL-selected cohort also lacked numerous pandemic pathogenic STs such as ST12, ST58, ST73, ST95, ST117, and ST127 (Manges et al., 2019). Consequently, such methodological approaches risk underrepresenting or omitting non-ESBL-expressing pathogenic *E. coli* strains, potentially leading to a distorted perception of the microbial ecology and the actual spectrum of AMR within a population.

4.2. Most ESBL-selected isolates were MDR while most non-selected isolates carried no acquired ARGs

We observed a significant discrepancy in AMR profiles between ESBL-selected and non-selected isolates. Predominantly, ESBL-selected isolates were MDR, harbouring an average of six ARGs. These genes predominantly conferred resistance to extended-spectrum β-lactams (93 %, with aforementioned ST410 isolates as notable exceptions), tetracycline (59 %), streptomycin (59 %), macrolides (58 %), sulphonamides (58 %), and 41 % exhibited dual *gyrA/parC* mutations indicating fluoroquinolone resistance. This AMR pattern contrasts starkly with that of non-selected isolates, where 80 % exhibited no transferable ARGs, only 1 % had dual *gyrA/parC* mutations, and none possessed the $bla_{\text{CTX-M}}$ genes. There is currently a lack of genomic studies directly comparing ARGs in ESBL-selected and non-ESBL-selected *E. coli* in wastewater environments. However, similar trends have been noted in other contexts. For example, in a study comparing ESBL-producing and non-ESBL-producing *E. coli* isolates causing bacteraemia the average ARG count was 7 for the former and 1 for the latter (van Hout et al., 2020).

Globally, $bla_{\text{CTX-M-15}}$ and $bla_{\text{CTX-M-14}}$ are the most prevalent ESBL genes, with some geographic and host/environmental variability (Bevan et al., 2017). In our study, $bla_{\text{CTX-M-15}}$ was the most common in the ESBL-selected group, found across 16 STs, accounting for 54 % of the ESBL genes. Notably, $bla_{\text{CTX-M-27}}$, a single nucleotide polymorphism variant of $bla_{\text{CTX-M-14}}$ and more resistant to cefotaxime, is increasingly reported to be outcompeting other variants (Bevan et al., 2017). Here, $bla_{\text{CTX-M-14}}$ was only identified in two isolates, while $bla_{\text{CTX-M-27}}$ was the second most common ESBL gene (33 %) and was the dominate ESBL gene in ST131 isolates. This dominance is due to most ST131 isolates belonging to clade A, previously noted to carry $bla_{\text{CTX-M-27}}$ in Australia (Li et al., 2021), rather than clade C2 (75% vs 11 %). Interestingly, we found a near equal distribution of $bla_{\text{CTX-M}}$ genes on IncF plasmids and chromosomal locations, with all chromosomal integrations appearing to be IS*Ecp1*-mediated. This supports growing evidence of chromosomally integrated $bla_{\text{CTX-M}}$ genes (Biggel et al., 2023; Li et al., 2021; Shawa et al., 2021; Zahra et al., 2018), a concerning trend that suggests increasing propensity for vertical inheritance and stable maintenance of these genes, independent of antibiotic usage.

Significant for AMR surveillance, our data indicates that selection for ESBL-producing isolates also enables surveillance of other important ARGs. Nevertheless, it is noteworthy that certain ARGs of high clinical importance, specifically *tet(X4)* and $bla_{\text{CMY-2}}$ (AmpC ESBL), were exclusively detected in the non-selected isolates. In particular, mobile *tet(X4)* genes, first reported in 2019, have attracted considerable international concern because they inactivate all tetracyclines, including the last-resort antibiotic tigecycline (He et al., 2019). These data highlight the complexity and variability of AMR in *E. coli*, emphasising the importance of comprehensive surveillance strategies that include both ESBL-selected and non-selected isolates to accurately assess the AMR spectrum within a population.

4.3. ExPEC is the dominate pathotype in both ESBL-selected and non-selected cohorts

ExPEC are globally recognized as the primary causative agents of UTIs and BSIs. The presence of ExPEC in wastewater, if not adequately treated, poses a public health risk, particularly if recycled water is then used for recreational purposes or agricultural irrigation. Alarmingly, several studies have indicated that *E. coli* strains surviving wastewater treatment are predominately ExPEC (Anastasi et al., 2013; Calhau et al., 2014; Paulshus et al., 2019). Here we found that ExPEC was the dominate pathotype in both ESBL-selected and non-selected cohorts (60 % and 23 %, respectively). The higher levels of ExPEC in ESBL-selected isolates aligns with the high numbers of ST131 in this cohort, but the portion of ExPEC in non-selected wastewater isolates mirrors findings from Canada, where the rate is 24 % (Frigon et al., 2013). Notably, we uncovered a low prevalence of intestinal pathogenic *E. coli* (IPEC) in either group (4 % of total isolates), and no Shiga toxin-producing *E. coli* (STEC). This aligns with the infrequent occurrence of STEC, and rare instances of haemolytic uremic syndrome (HUS) (associated STEC disease) cases reported in Australia, which are approximately 0.3 cases per 100,000 population in NSW (The OzFoodNet Working Group, 2021).

4.4. The tLST was present in 20 % of non-selected isolates, and absent from ESBL-selected isolates

The tLST is a genomic island flanked by transposons that confers tolerance to elevated temperature, pressure, and chlorine (Kamal et al., 2021). This locus is particularly relevant in contexts where thermal and chlorination methods are employed as antimicrobial measures, such as in meat and milk processing (heat) and wastewater sanitation (chlorination). These practices could inadvertently promote the selection of tLST-positive bacteria. A recent comprehensive study analysing 18,959 *E. coli* genomes made several key observations about tLST distribution (Zhang and Yang, 2022), namely that the tLST is found in 2.7 % of *E. coli* and only in phylogroup A, B1 and C (thus absent in most pathogen-associated phylogroups such as B2, D and E), that it was more prevalent in wastewater and raw milk isolates, that its more commonly chromosomally located but also found on plasmids, and that its presence was negatively correlated with the presence of ARGs and VAGs (Zhang and Yang, 2022). While our observations were in agreement with this study, a recent study on *E. coli* from silver gulls from Australia found a considerable portion of tLST+ isolates also carried critically important ARGs, including carbapenemase bla_{IMP} genes (Wyrsch et al., 2024). Among our 40 tLST+ isolates all were either phylogroup A (85 %) or B1 (15%), the locus was predominately located in the chromosome (86 %) but also present on some novel putative plasmids, and that they carried almost no acquired ARGs (one isolate carried one *qnrB4*). However, while most isolates carried no major VAGs, two tLST+ ST399 isolates had HPI, which is unusual as this pathogenicity island is ExPEC-associated (Galardini et al., 2020) and previous studies have reported a lack of VAGs in this STs (Behruznia and Gordon, 2022). ST399 warrants close monitoring as, though thought previously to be wastewater specific, it was recently reported to have captured a plasmid encoding the carbapenemase $bla_{\text{OXA-48}}$ and caused a hospital outbreak in the UK (Ledda et al., 2020). We also observed that tLST+ isolates carried significantly more metal resistance genes, particularly those conferring resistance to arsenic, copper, mercury, nickel, and silver. A considerable amount of these metals from both anthropogenic and natural sources end up in wastewater (Martin et al., 2021; Yuan et al., 2019), which may further contribute to selection pressure favouring tLST+ strains.

4.5. Limitations

This study acknowledges several limitations in the methodologies used, namely the reliance on culture-based short-read WGS. Short-read sequencing technologies have inherent limitations in resolving repeat-rich structures, such as plasmids. Therefore, conclusions regarding the putative plasmids carrying tLST should be interpreted with caution. Future studies could benefit from integrating long-read sequencing technologies, which offer improved resolution of plasmid structures and their genetic contexts.

While culture-based WGS allows for detailed characterisation of individual isolates, including precise identification of ARGs, VAGs, and STs, and a deep understanding of phylogeny, it introduces selection biases. Culturing *E. coli* may preferentially select strains that thrive under laboratory conditions. In contrast, metagenomics enables comprehensive analysis of the dominant species representing a greater cross section of the of the entire microbial community within detection limits, including all abundant *E. coli* lineages present within a sample. However, metagenomics is limited by its lower resolution and specificity, particularly in attributing specific genes to species. Given that this

study aimed to investigate *E. coli* specifically, culture-based WGS—offering higher resolution, specificity, and accuracy—was deemed most suited for providing detailed insights into the genomic characteristics and of wastewater *E. coli*.

Recognising these limitations, future studies should aim to incorporate both short-read and long-read sequencing technologies, along with culture-independent metagenomic approaches, to achieve a more comprehensive understanding of plasmid dynamics and role *E. coli* has in the broader microbial community in wastewater ecosystems.

5. Conclusion

This comprehensive genomic analysis of both ESBL-selected and non-selected *E. coli* isolates from wastewater in NSW, Australia, offers valuable insights into the diversity, AMR, and pathogenicity traits within this population. While the expansion of genomic surveillance may introduce additional costs, our study highlights the importance of incorporating a broad spectrum of isolates to capture the full extent of genomic diversity and AMR patterns, which ultimately enhances public health protection and facilitates a better understanding of AMR dynamics. The predominance of certain sequence types, such as ST131 in ESBL-selected isolates, underscores the potential for selection bias inherent in AMR surveillance methodologies. Additionally, the presence of the tLST in a significant portion of non-selected isolates, and its association with metal resistance genes, points to the environmental pressures shaping bacterial populations in wastewater ecosystems. Our findings emphasise the need for vigilant monitoring of *E. coli* populations, particularly in environments impacted by human activities, to better understand and mitigate the public health risks posed by AMR and pathogenic *E. coli* strains. Such efforts are crucial for guiding public health policies and informing strategies to combat the escalating challenge of AMR. Overall, our research contributes novel insight into the growing body of knowledge on *E. coli* ecology and AMR, reinforcing the significance of comprehensive, unbiased genomic surveillance in understanding and managing these critical public health threats.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2024.175079.

Funding Statement

This project was supported by an Australian Government Medical Research Future Fund Grant (MRFF75873): Disruptive Technologies to Trace, Track & Tackle Antibiotic Resistance, and from the Australian Centre for Genomic Epidemiological Microbiology (AusGEM), a collaborative partnership between the NSW Department of Primary Industries and the University of Technology Sydney.

CRediT authorship contribution statement

Veronica M. Jarocki: Writing – review & editing, Writing – original draft, Visualization, Investigation, Formal analysis, Conceptualization. **Dmitriy Li:** Writing – review & editing, Formal analysis. **Daniel R. Bogema:** Writing – review & editing, Resources, Investigation. **Jerald Yam:** Investigation, Data curation. **Cheryl Jenkins:** Writing – review & editing, Supervision, Resources, Project administration, Methodology. **Faisal I. Hai:** Writing – review & editing, Resources, Investigation. **Steven P. Djordjevic:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Raw reads from all 281 isolates described in this study have been deposited in NCBI under BioProject number PRJNA1087400

References

Alcock, B.P., Raphenya, A.R., Lau, T.T.Y., Tsang, K.K., Bouchard, M., Edalatmand, A., Huynh, W., Nguyen, A.-L.V., Cheng, A.A., Liu, S., Min, S.Y., Miroshnichenko, A., Tran, H.-K., Werfalli, R.E., Nasir, J.A., Oloni, M., Speicher, D.J., Florescu, A., Singh, B., Faltyn, M., Hernandez-Koutoucheva, A., Sharma, A.N., Bordeleau, E., Pawlowski, A.C., Zubyk, H.L., Dooley, D., Griffiths, E., Maguire, F., Winsor, G.L., Beiko, R.G., Brinkman, F.S.L., Hsiao, W.W.L., Domselaar, G.V., McArthur, A.G., 2020. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res. 48, D517–D525. https://doi.org/10.1093/nar/gkz935.

Anastasi, E.M., Wohlsen, T.D., Stratton, H.M., Katouli, M., 2013. Survival of Escherichia coli in two sewage treatment plants using UV irradiation and chlorination for disinfection. Water Res. 47, 6670–6679. https://doi.org/10.1016/j.watres.2013.09.008.

Arndt, D., Grant, J.R., Marcu, A., Sajed, T., Pon, A., Liang, Y., Wishart, D.S., 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res. 44, W16–W21. https://doi.org/10.1093/nar/gkw387.

Arredondo-Alonso, S., Rogers, M.R.C., Braat, J.C., Verschuuren, T.D., Top, J., Corander, J., Willems, R.J.L., Schürch, A.C.Y., 2018. Mlplasmids: a user-friendly tool to predict plasmid- and chromosome-derived sequences for single species. Microbial. Genomics 4, e000224. https://doi.org/10.1099/mgen.0.000224.

Atterby, C., Börjesson, S., Ny, S., Järhult, J.D., Byfors, S., Bonnedahl, J., 2017. ESBL-producing Escherichia coli in Swedish gulls—a case of environmental pollution from humans? PLoS One 12, e0190380. https://doi.org/10.1371/journal.pone.0190380.

AURA, 2023. AURA 2023: Fifth Australian Report on Antimicrobial Use and Resistance in Human Health.

Baker, K.S., Jauneikaite, E., Nunn, J.G., Midega, J.T., Atun, R., Holt, K.E., Walia, K., Howden, B.P., Tate, H., Okeke, I.N., Carattoli, A., Hsu, L.Y., Hopkins, K.L., Muloi, D. M., Wheeler, N.E., Aanensen, D.M., Mason, L.C.E., Rodgus, J., Hendriksen, R.S., Essack, S.Y., Egyir, B., Halpin, A.L., MacCannell, D.R., Campos, J., Srikantiah, P., Feasey, N.A., Peacock, S.J., 2023. Evidence review and recommendations for the implementation of genomics for antimicrobial resistance surveillance: reports from an international expert group. The Lancet Microbe 4, e1035–e1039. https://doi.org/10.1016/S2666-5247(23)00281-1.

Behruznia, M., Gordon, D.M., 2022. Molecular and metabolic characteristics of wastewater associated Escherichia coli strains. Environ. Microbiol. Rep. 14, 646–654. https://doi.org/10.1111/1758-2229.13076.

Bevan, E.R., Jones, A.M., Hawkey, P.M., 2017. Global epidemiology of CTX-M β-lactamases: temporal and geographical shifts in genotype. J. Antimicrob. Chemother. 72, 2145–2155. https://doi.org/10.1093/jac/dkx146.

Biggel, M., Hoehn, S., Frei, A., Dassler, K., Jans, C., Stephan, R., 2023. Dissemination of ESBL-producing E. Coli ST131 through wastewater and environmental water in Switzerland. Environ. Pollut. 337, 122476 https://doi.org/10.1016/j.envpol.2023.122476.

Blanco, J., Mora, A., Mamani, R., López, C., Blanco, M., Dahbi, G., Herrera, A., Marzoa, J., Fernández, V., de la Cruz, F., Martínez-Martínez, L., Alonso, M.P., Nicolas-Chanoine, M.-H., Johnson, J.R., Johnston, B., López-Cerero, L., Pascual, A., Rodríguez-Baño, J., Spanish Group for Nosocomial Infections (GEIH), 2013. Four main virotypes among extended-spectrum-β-lactamase-producing isolates of Escherichia coli O25b:H4-B2-ST131: bacterial, epidemiological, and clinical characteristics. J. Clin. Microbiol. 51, 3358–3367. https://doi.org/10.1128/JCM.01555-13.

Bréchet, C., Plantin, J., Sauget, M., Thouverez, M., Talon, D., Cholley, P., Guyeux, C., Hocquet, D., Bertrand, X., 2014. Wastewater treatment plants release large amounts of extended-spectrum β-lactamase–producing Escherichia coli into the environment. Clin. Infect. Dis. 58 https://doi.org/10.1093/cid/ciu190.

Calhau, V., Mendes, C., Pena, A., Mendonça, N., Da Silva, G.J., 2014. Virulence and plasmidic resistance determinants of Escherichia coli isolated from municipal and hospital wastewater treatment plants. J. Water Health 13, 311–318. https://doi.org/10.2166/wh.2014.327.

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.

Clermont, O., Christenson, J.K., Denamur, E., Gordon, D.M., 2013. The Clermont Escherichia coli phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. Environ. Microbiol. Rep. 5, 58–65. https://doi.org/10.1111/1758-2229.12019.

Conforti, S., Holschneider, A., Sylvestre, E., Julian, T.R., 2024. Monitoring ESBL-Escherichia coli in Swiss wastewater between November 2021 and November 2022: insights into population carriage. mSphere 9 (5). https://doi.org/10.1128/msphere.00760-23.

Cormier, A.C., Chalmers, G., Zamudio, R., Mulvey, M.R., Mather, A.E., Boerlin, P., 2022. Diversity of blaCTX-M-1-carrying plasmids recovered from Escherichia coli isolated from Canadian domestic animals. PLoS One 17 (3), e0264439. https://doi.org/10.1371/journal.pone.0264439.

Croucher, N.J., Page, A.J., Connor, T.R., Delaney, A.J., Keane, J.A., Bentley, S.D., Parkhill, J., Harris, S.R., 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.
Cummins, M.L., Reid, C.J., Djordjevic, S.P., 2022. F plasmid lineages in escherichia coli st95: implications for host range, antibiotic resistance, and zoonoses. mSystems. https://doi.org/10.1128/msystems.01212-21 e01212-21.
Darling, A.C.E., Mau, B., Blattner, F.R., Perna, N.T., 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. 14, 1394–1403. https://doi.org/10.1101/gr.2289704.
Davidova-Gerzova, L., Lausova, J., Sukkar, I., Nesporova, K., Nechutna, L., Vlkova, K., Chudejova, K., Krutova, M., Palkovicova, J., Kaspar, J., Dolejska, M., 2023. Hospital and community wastewater as a source of multidrug-resistant ESBL-producing Escherichia coli. Front. Cell. Infect. Microbiol. 13.
Day, M.J., Rodríguez, I., van Essen-Zandbergen, A., Dierikx, C., Kadlec, K., Schink, A., Wu, G., Chattaway, M.A., DoNascimento, V., Wain, J., Helmuth, R., Guerra, B., Schwarz, S., Threlfall, J., Woodward, M.J., Coldham, N., Mevius, D., Woodford, N., 2016. Diversity of STs, plasmids and ESBL genes among Escherichia coli from humans, animals and food in Germany, the Netherlands and the UK. J. Antimicrob. Chemother. 71 (5), 1178–1182. https://doi.org/10.1093/jac/dkv485.
Denamur, E., Clermont, O., Bonacorsi, S., Gordon, D., 2021. The population genetics of pathogenic Escherichia coli. Nat. Rev. Microbiol. 19, 37–54. https://doi.org/10.1038/s41579-020-0416-x.
Djordjevic, S.P., Jarocki, V.M., Seemann, T., Cummins, M.L., Watt, A.E., Drigo, B., Wyrsch, E.R., Reid, C.J., Donner, E., Howden, B.P., 2023. Genomic surveillance for antimicrobial resistance — a one health perspective. Nat. Rev. Genet. 1–16 https://doi.org/10.1038/s41576-023-00649-y.
Dolejska, M., Duskova, E., Rybarikova, J., Janoszowska, D., Roubalova, E., Dibdakova, K., Maceckova, G., Kohoutova, L., Literak, I., Smola, J., Cizek, A., 2011. Plasmids carrying blaCTX-M-1 and qnr genes in Escherichia coli isolates from an equine clinic and a horseback riding Centre. J. Antimicrob. Chemother. 66, 757–764. https://doi.org/10.1093/jac/dkq500.
Finn, T.J., Scriver, L., Lam, L., Duong, M., Peirano, G., Lynch, T., Dong, T., Pitout, J.D.D., DeVinney, R., 2020. A comprehensive account of Escherichia coli sequence type 131 in wastewater reveals an abundance of fluoroquinolone-resistant clade a strains. Appl. Environ. Microbiol. 86, e01913–e01919. https://doi.org/10.1128/AEM.01913-19.
Florensa, A.F., Kaas, R.S., Clausen, P.T.L.C., Aytan-Aktug, D., Aarestrup, F.M., 2022. ResFinder – an open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes. Microb. Genom. 8, 000748 https://doi.org/10.1099/mgen.0.000748.
Frigon, D., Biswal, B.K., Mazza, A., Masson, L., Gehr, R., 2013. Biological and physicochemical wastewater treatment processes reduce the prevalence of virulent Escherichia coli. Appl. Environ. Microbiol. 79, 835–844. https://doi.org/10.1128/AEM.02789-12.
Gaio, D., Anantanawat, K., To, J., Liu, M., Monahan, L., Darling, A.E., 2022. Hackflex: low-cost, high-throughput, Illumina Nextera flex library construction. Microb. Genom. 8 https://doi.org/10.1099/mgen.0.000744.
Galardini, M., Clermont, O., Baron, A., Busby, B., Dion, S., Schubert, S., Beltrao, P., Denamur, E., 2020. Major role of iron uptake systems in the intrinsic extra-intestinal virulence of the genus Escherichia revealed by a genome-wide association study. PLoS Genet. 16, e1009065 https://doi.org/10.1371/journal.pgen.1009065.
Gillings, M.R., Gaze, W.H., Pruden, A., Smalla, K., Tiedje, J.M., Zhu, Y.-G., 2015. Using the class 1 integron-integrase gene as a proxy for anthropogenic pollution. ISME J. 9, 1269–1279. https://doi.org/10.1038/ismej.2014.226.
He, T., Wang, R., Liu, D., Walsh, T.R., Zhang, R., Lv, Y., Ke, Y., Ji, Q., Wei, R., Liu, Z., Shen, Y., Wang, G., Sun, L., Lei, L., Lv, Z., Li, Y., Pang, M., Wang, L., Sun, Q., Fu, Y., Song, H., Hao, Y., Shen, Z., Wang, S., Chen, G., Wu, C., Shen, J., Wang, Y., 2019. Emergence of plasmid-mediated high-level tigecycline resistance genes in animals and humans. Nat. Microbiol. 4, 1450–1456. https://doi.org/10.1038/s41564-019-0445-2.
Hendriksen, R.S., Munk, P., Njage, P., van Bunnik, B., McNally, L., Lukjancenko, O., Röder, T., Nieuwenhuijse, D., Pedersen, S.K., Kjeldgaard, J., Kaas, R.S., Clausen, P.T. L.C., Vogt, J.K., Leekitcharoenphon, P., van de Schans, M.G.M., Zuidema, T., de Roda Husman, A.M., Rasmussen, S., Petersen, B., Amid, C., Cochrane, G., Sicheritz-Ponten, T., Schmitt, H., Alvarez, J.R.M., Aidara-Kane, A., Pamp, S.J., Lund, O., Hald, T., Woolhouse, M., Koopmans, M.P., Vigre, H., Petersen, T.N., Aarestrup, F.M., 2019. Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. Nat. Commun. 10, 1124. https://doi.org/10.1038/s41467-019-08853-3.
Jang, J., Hur, H.-G., Sadowsky, M.J., Byappanahalli, M.N., Yan, T., Ishii, S., 2017. Environmental Escherichia coli: ecology and public health implications—a review. J. Appl. Microbiol. 123, 570–581. https://doi.org/10.1111/jam.13468.
Jarocki, V.M., Heß, S., Anantanawat, K., Berendonk, T.U., Djordjevic, S.P., 2021. Multidrug-resistant lineage of Enterotoxigenic Escherichia coli ST182 with serotype O169:H41 in airline waste. Front. Microbiol. 12, 2301. https://doi.org/10.3389/fmicb.2021.731050.
Joensen, K.G., Tetzschner, A.M.M., Iguchi, A., Aarestrup, F.M., Scheutz, F., 2015. Rapid and easy in silico serotyping of Escherichia coli isolates by use of whole-genome sequencing data. J. Clin. Microbiol. 53, 2410–2426. https://doi.org/10.1128/JCM.00008-15.
Johnson, J.R., Stell, A.L., 2000. Extended virulence genotypes of Escherichia coli strains from patients with Urosepsis in relation to phylogeny and host compromise. J. Infect. Dis. 181, 261–272. https://doi.org/10.1086/315217.
Johnson, J.R., Porter, S., Johnston, B., Kuskowski, M.A., Spurbeck, R.R., Mobley, H.L.T., Williamson, D.A., 2015. Host characteristics and bacterial traits predict experimental virulence for Escherichia coli bloodstream isolates from patients with Urosepsis. Open forum. Infect. Dis. 2, ofv083. https://doi.org/10.1093/ofid/ofv083.
Johnson, T.J., Nolan, L.K., 2009. Pathogenomics of the virulence plasmids of Escherichia coli. Microbiol. Mol. Biol. Rev. 73, 750–774. https://doi.org/10.1128/MMBR.00015-09.
Johnson, T.J., Danzeisen, J.L., Youmans, B., Case, K., Llop, K., Munoz-Aguayo, J., Flores-Figueroa, C., Aziz, M., Stoesser, N., Sokurenko, E., Price, L.B., Johnson, J.R., 2016. Separate F-type plasmids have shaped the evolution of the H30 Subclone of *Escherichia coli* Sequence Type 131. mSphere 1. https://doi.org/10.1128/mSphere.00121-16 e00121-16.
Jørgensen, S.B., Søraas, A.V., Arnesen, L.S., Leegaard, T.M., Sundsfjord, A., Jenum, P.A., 2017. A comparison of extended spectrum β-lactamase producing Escherichia coli from clinical, recreational water and wastewater samples associated in time and location. PLoS One 12, e0186576. https://doi.org/10.1371/journal.pone.0186576.
Kallonen, T., Brodrick, H.J., Harris, S.R., Corander, J., Brown, N.M., Martin, V., Peacock, S.J., Parkhill, J., 2017. Systematic longitudinal survey of invasive Escherichia coli in England demonstrates a stable population structure only transiently disturbed by the emergence of ST131. Genome Res. 27, 1437–1449. https://doi.org/10.1101/gr.216606.116.
Kamal, S.M., Simpson, D.J., Wang, Z., Gänzle, M., Römling, U., 2021. Horizontal transmission of stress resistance genes shape the ecology of Beta- and Gamma-Proteobacteria. Front. Microbiol. 12, 696522 https://doi.org/10.3389/fmicb.2021.696522.
Kudinha, T., Kong, F., 2022. Possible step-up in prevalence for Escherichia coli ST131 from fecal to clinical isolates: inferred virulence potential comparative studies within phylogenetic group B2. J. Biomed. Sci. 29, 78. https://doi.org/10.1186/s12929-022-00862-7.
Larsen, M.V., Cosentino, S., Rasmussen, S., Friis, C., Hasman, H., Marvig, R.L., Jelsbak, L., Sicheritz-Pontén, T., Ussery, D.W., Aarestrup, F.M., Lund, O., 2012. Multilocus sequence typing of total-genome-sequenced bacteria. J. Clin. Microbiol. 50, 1355–1361. https://doi.org/10.1128/JCM.06094-11.
Leclerc, H., Mossel, D.A.A., Edberg, S.C., Struijk, C.B., 2001. Advances in the bacteriology of the coliform group: their suitability as markers of microbial water safety. Ann. Rev. Microbiol. 55, 201–234. https://doi.org/10.1146/annurev.micro.55.1.201.
Ledda, A., Cummins, M., Shaw, L.P., Jauneikaite, E., Cole, K., Lasalle, F., Barry, D., Rosmarin, C., Anaraki, S., Wareham, D., Stoesser, N., Paul, J., Manuel, R., Cherian, B. P., Didelot, X., 2020. Hospital outbreak of carbapenem-resistant Enterobacteriales associated with an OXA-48 plasmid carried mostly by Escherichia coli ST399. bioRxiv. https://doi.org/10.1101/2020.06.15.148189.
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.
Li, D., Wyrsch, E.R., Elankumaran, P., Dolejska, M., Marenda, M.S., Browning, G.F., Bushell, R.N., McKinnon, J., Chowdhury, P.R., Hitchick, N., Miller, N., Donner, E., Drigo, B., Baker, D., Charles, I.G., Kudinha, T., Jarocki, V.M., Djordjevic, S.P., 2021. Genomic comparisons of Escherichia coli ST131 from Australia. Microbial. Genomics 7. https://doi.org/10.1099/mgen.0.000721, 000721.
Li, D., Elankumaran, P., Kudinha, T., Kidsley, A.K., Trott, D.J., Jarocki, V.M., Djordjevic, S.P., 2023. Dominance of Escherichia coli sequence types ST73, ST95, ST127 and ST131 in Australian urine isolates: a genomic analysis of antimicrobial resistance and virulence linked to F plasmids. Microb. Genom. 9, mgen001068. https://doi.org/10.1099/mgen.0.001068.
Liu, B., Zheng, D., Jin, Q., Chen, L., Yang, J., 2019. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. Nucleic Acids Res. 47, D687–D692. https://doi.org/10.1093/nar/gky1080.
Manges, A.R., Geum, H.M., Guo, A., Edens, T.J., Fibke, C.D., Pitout, J.D.D., 2019. Global Extrain- testinal pathogenic Escherichia coli (ExPEC) lineages. Clin. Microbiol. Rev. 32 https://doi.org/10.1128/CMR.00135-18.
Martin, C., Stebbins, B., Ajmani, A., Comendul, A., Hamner, S., Hasan, N.A., Colwell, R., Ford, T., 2021. Nanopore-based metagenomics analysis reveals prevalence of mobile antibiotic and heavy metal resistome in wastewater. Ecotoxicology 30, 1572–1585. https://doi.org/10.1007/s10646-020-02342-w.
Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., Lanfear, R., 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. 37, 1530–1534. https://doi.org/10.1093/molbev/msaa015.
Murray, C.J., Ikuta, K.S., Sharara, F., Swetschinski, L., Aguilar, G.R., Gray, A., Han, C., Bisignano, C., Rao, P., Wool, E., Johnson, S.C., Browne, A.J., Chipeta, M.G., Fell, F., Hackett, S., Haines-Woodhouse, G., Hamadani, B.H.K., Kumaran, E.A.P., McManigal, B., Agarwal, R., Akech, S., Albertson, S., Amuasi, J., Andrews, J., Aravkin, A., Ashley, E., Bailey, F., Baker, S., Basnyat, B., Bekker, A., Bender, R., Bethou, A., Bielicki, J., Boonkasidecha, S., Bukosia, J., Carvalheiro, C., Castañeda-Orjuela, C., Chansamouth, V., Chaurasia, S., Chiurchiù, S., Chowdhury, F., Cook, A. J., Cooper, B., Cressey, T.R., Criollo-Mora, E., Cunningham, M., Darboe, S., Day, N.P. J., Luca, M.D., Dokova, K., Dramowski, A., Dunachie, S.J., Eckmanns, T., Eibach, D., Emami, A., Feasey, N., Fisher-Pearson, N., Forrest, K., Garrett, D., Gastmeier, P., Giref, A.Z., Greer, R.C., Gupta, V., Haller, S., Haselbeck, A., Hay, S.I., Holm, M., Hopkins, S., Iregbu, K.C., Jacobs, J., Jarovsky, D., Javanmardi, F., Khorana, M., Kissoon, N., Kobeissi, E., Kostyanev, T., Krapp, F., Krumkamp, R., Kumar, A., Kyu, H. H., Lim, C., Limmathurotsakul, D., Loftus, M.J., Lunn, M., Ma, J., Mturi, N., Munera-Huertas, T., Musicha, P., Mussi-Pinhata, M.M., Nakamura, T., Nanavati, R., Nangia, S., Newton, P., Ngoun, C., Novotney, A., Nwakanma, D., Obiero, C.W., Olivas-Martinez, A., Olliaro, P., Ooko, E., Ortiz-Brizuela, E., Peleg, A.Y., Perrone, C., Plakkal, N., Ponce-de-Leon, A., Raad, M., Ramdin, T., Riddell, A., Roberts, T., Robotham, J.V., Roca, A., Rudd, K.E., Russell, N., Schnall, J., Scott, J.A.G., Shivamallappa, M., Sifuentes-Osornio, J., Steenkeste, N., Stewardson, A.J., Stoeva, T., Tasak, N., Thaiprakong, A., Thwaites, G., Turner, C., Turner, P., van

Doorn, H.R., Velaphi, S., Vongpradith, A., Vu, H., Walsh, T., Waner, S., Wangrangsimakul, T., Wozniak, T., Zheng, P., Sartorius, B., Lopez, A.D., Stergachis, A., Moore, C., Dolecek, C., Naghavi, M., 2022. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. Lancet 399, 629–655. https://doi.org/10.1016/S0140-6736(21)02724-0.

Osińska, A., Korzeniewska, E., Harnisz, M., Niestępski, S., 2017. The prevalence and characterization of antibiotic-resistant and virulent Escherichia coli strains in the municipal wastewater system and their environmental fate. Sci. Total Environ. 577, 367–375. https://doi.org/10.1016/j.scitotenv.2016.10.203.

Page, A.J., Taylor, B., Delaney, A.J., Soares, J., Seemann, T., Keane, J.A., Harris, S.R., 2016. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. Microb. Genom. 2, e000056 https://doi.org/10.1099/mgen.0.000056.

Paul, D., Babenko, D., Toleman, M.A., 2020. Human carriage of cefotaxime-resistant Escherichia coli in north-East India: an analysis of STs and associated resistance mechanisms. J. Antimicrob. Chemother. 75, 72–76. https://doi.org/10.1093/jac/dkz416.

Paulshus, E., Thorell, K., Guzman-Otazo, J., Joffre, E., Colque, P., Kühn, I., Möllby, R., Sørum, H., Sjöling, Å., 2019. Repeated isolation of extended-Spectrum-β-lactamase-positive Escherichia coli sequence types 648 and 131 from community wastewater indicates that sewage systems are important sources of emerging clones of antibiotic-resistant Bacteria. Antimicrob. Agents Chemother. 63 https://doi.org/10.1128/AAC.00823-19.

Pitout, J.D.D., Chen, L., 2023. The significance of epidemic plasmids in the success of multidrug resistant drug pandemic extraintestinal pathogenic Escherichia coli. Infect. Dis. Ther. 12, 1029–1041. https://doi.org/10.1007/s40121-023-00791-4.

Pitout, J.D.D., Laupland, K.B., 2008. Extended-spectrum beta-lactamase-producing Entero- bacteriaceae: an emerging public-health concern. Lancet Infect. Dis. 8, 159–166. https://doi.org/10.1016/S1473-3099(08)70041-0.

Price, M.N., Dehal, P.S., Arkin, A.P., 2010. FastTree 2 – approximately maximum-likelihood trees for large alignments. PLoS One 5, e9490. https://doi.org/10.1371/journal.pone.0009490.

Raven, K.E., Ludden, C., Gouliouris, T., Blane, B., Naydenova, P., Brown, N.M., Parkhill, J., Peacock, S.J., 2019. Genomic surveillance of Escherichia coli in municipal wastewater treatment plants as an indicator of clinically relevant pathogens and their resistance genes. Microb. Genom. 5, e000267 https://doi.org/10.1099/mgen.0.000267.

Reid, C.J., Cummins, M.L., Börjesson, S., Brouwer, M.S.M., Hasman, H., Hammerum, A. M., Roer, L., Hess, S., Berendonk, T., Nešporová, K., Haenni, M., Madec, J.-Y., Bethe, A., Michael, G.B., Schink, A.-K., Schwarz, S., Dolejska, M., Djordjevic, S.P., 2022. A role for ColV plasmids in the evolution of pathogenic Escherichia coli ST58. Nat. Commun. 13, 1–15. https://doi.org/10.1038/s41467-022-28342-4.

Robertson, J., Nash, J.H.E., 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.

Roer, L., Tchesnokova, V., Allesøe, R., Muradova, M., Chattopadhyay, S., Ahrenfeldt, J., Thomsen, M.C.F., Lund, O., Hansen, F., Hammerum, A.M., Sokurenko, E., Hasman, H., 2017. Develop- ment 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.

Roer, L., Overballe-Petersen, S., Hansen, F., Schønning, K., Wang, M., Røder, B.L., Hansen, D.S., Justesen, U.S., Andersen, L.P., Fulgsang-Damgaard, D., Hopkins, K.L., Woodford, N., Falgenhauer, L., Chakraborty, T., Samuelsen, Ø., Sjöström, K., Johannesen, T.B., Ng, K., Nielsen, J., Ethelberg, S., Stegger, M., Hammerum, A.M., Hasman, H., 2018. *Escherichia coli* sequence type 410 is causing new international high-risk clones. mSphere 3. https://doi.org/10.1128/msphere.00337-18.

Royer, G., Clermont, O., Marin, J., Condamine, B., Dion, S., Blanquart, F., Galardini, M., Denamur, E., 2023. Epistatic interactions between the high pathogenicity island and other iron uptake systems shape Escherichia coli extra-intestinal virulence. Nat. Commun. 14, 3667. https://doi.org/10.1038/s41467-023-39428-y.

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

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

Sekizuka, T., Tanaka, R., Hashino, M., Yatsu, K., Kuroda, M., 2022. Comprehensive genome and Plasmidome analysis of antimicrobial resistant Bacteria in wastewater treatment plant effluent of Tokyo. Antibiotics 11, 1283. https://doi.org/10.3390/antibiotics11101283.

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.

Sherry, N.L., Horan, K., Ballard, S.A., da Silva, A.G., Gorrie, C.L., Schultz, M.B., Stevens, K., Valcanis, M., Sait, M., Stinear, T.P., Howden, B.P., Seemann, T., 2022. Bridging the gap between bioinformatics and the clinical and public health microbiology laboratory: An ISO-accredited genomics workflow for antimicrobial resistance. https://doi.org/10.1101/2022.05.26.493552.

Shibuki, R., Nishiyama, M., Mori, M., Baba, H., Kanamori, H., Watanabe, T., 2023. Characterization of extended-spectrum β-lactamase-producing Escherichia coli isolated from municipal and hospital wastewater in Japan. J. Glob. Antimicrob. Resist. 32, 145–151. https://doi.org/10.1016/j.jgar.2023.02.002.

Sullivan, M.J., Petty, N.K., Beatson, S.A., 2011. Easyfig: a genome comparison visualizer. Bioinformatics 27, 1009–1010. https://doi.org/10.1093/bioinformatics/btr039.

Sun, J., Chen, C., Cui, C.-Y., Zhang, Y., Liu, X., Cui, Z.-H., Ma, X.-Y., Feng, Y., Fang, L.-X., Lian, X.-L., Zhang, R.-M., Tang, Y.-Z., Zhang, K.-X., Liu, H.-M., Zhuang, Z.-H., Zhou, S.-D., Lv, J.-N., Du, H., Huang, B., Yu, F.-Y., Mathema, B., Kreiswirth, B.N., Liao, X.-P., Chen, L., Liu, Y.-H., 2019. Plasmid-encoded tet(X) genes that confer high-level tigecycline resistance in Escherichia coli. Nat. Microbiol. 4, 1457–1464. https://doi.org/10.1038/s41564-019-0496-4.

Tanaka, H., Hayashi, W., Iimura, M., Taniguchi, Y., Soga, E., Matsuo, N., Kawamura, K., Arakawa, Y., Nagano, Y., Nagano, N., 2019. Wastewater as a probable environmental reservoir of extended-Spectrum-β-lactamase genes: detection of chimeric β-lactamases CTX-M-64 and CTX-M-123. Appl. Environ. Microbiol. 85 https://doi.org/10.1128/AEM.01740-19 e01740-19.

Tavares, R.D.S., Tacão, M., Figueiredo, A.S., Duarte, A.S., Esposito, F., Lincopan, N., Manaia, C.M., Henriques, I., 2020. Genotypic and phenotypic traits of blaCTX-M-carrying Escherichia coli strains from an UV-C-treated wastewater effluent. Water Res. 184, 116079 https://doi.org/10.1016/j.watres.2020.116079.

The OzFoodNet Working Group, 2021. Monitoring the incidence and causes of disease potentially transmitted by food in Australia: annual report of the OzFoodNet network, 2016. Commun. Dis. Intell. 45 https://doi.org/10.33321/cdi.2021.45.52.

Tonkin-Hill, G., MacAlasdair, N., Ruis, C., Weimann, A., Horesh, G., Lees, J.A., Gladstone, R.A., Lo, S., Beaudoin, C., Floto, R.A., Frost, S.D.W., Corander, J., Bentley, S.D., Parkhill, J., 2020. Producing polished prokaryotic pangenomes with the Panaroo pipeline. Genome Biol. 21, 180. https://doi.org/10.1186/s13059-020-02090-4.

Touchon, M., Perrin, A., de Sousa, J.A.M., Vangchhia, B., Burn, S., O'Brien, C.L., Denamur, E., Gordon, D., Rocha, E.P., 2020. Phylogenetic background and habitat drive the genetic diversification of Escherichia coli. PLoS Genet. 16, e1008866 https://doi.org/10.1371/journal.pgen.1008866.

Tu, J., Xue, T., Qi, K., Shao, Y., Huang, B., Wang, X., Zhou, X., 2016. The irp2 and fyuA genes in high Pathogenicity Islands are involved in the pathogenesis of infections caused by avian pathogenic Escherichia coli (APEC). Pol. J. Vet. Sci. 19, 21–29. https://doi.org/10.1515/pjvs-2016-0004.

van Hout, D., Verschuuren, T.D., Bruijning-Verhagen, P.C.J., Bosch, T., Schürch, A.C., Willems, R.J.L., Bonten, M.J.M., Kluytmans, J.A.J.W., 2020. Extended-spectrum beta-lactamase (ESBL)-producing and non-ESBL-producing *Escherichia coli* isolates causing bacteremia in the Netherlands (2014–2016) differ in clonal distribution, antimicrobial resistance gene and virulence gene content. PLoS One 15 (1), e0227604. https://doi.org/10.1371/journal.pone.0227604.

Wang, Z., Fang, Y., Zhi, S., Simpson, D.J., Gill, A., McMullen, L.M., Neumann, N.F., Gänzle, M.G., 2020. The locus of heat resistance confers resistance to chlorine and other oxidizing chemicals in Escherichia coli. Appl. Environ. Microbiol. 86 https://doi.org/10.1128/AEM.02123-19 e02123-19.

WHO, 2021. WHO integrated global surveillance on ESBL-producing *E. coli* using a "One Health" approach: implementation and opportunities.

Wyrsch, E.R., Nesporova, K., Tarabai, H., Jamborova, I., Bitar, I., Literak, I., Dolejska, M., Djordjevic, S.P., 2022. Urban Wildlife Crisis: Australian Silver Gull Is a Bystander Host to Widespread Clinical Antibiotic Resistance. mSystems 7, e0015822 https://doi.org/10.1128/msystems.00158-22.

Wyrsch, E.R., Hoye, B.J., Sanderson-Smith, M., Gorman, J., Maute, K., Cummins, M.L., Jarocki, V.M., Marenda, M.S., Dolejska, M., Djordjevic, S.P., 2024. The faecal microbiome of the Australian silver gull contains phylogenetically diverse ExPEC, aEPEC and Escherichia coli carrying the transmissible locus of stress tolerance. Sci. Total Environ. 1701815 https://doi.org/10.1016/j.scitotenv.2024.170815.

Yu, D., Ryu, K., Zhi, S., Otto, S.J.G., Neumann, N.F., 2022. Naturalized Escherichia coli in wastewater and the co-evolution of bacterial resistance to water treatment and antibiotics. Front. Microbiol. 13.

Yuan, L., Li, Z.-H., Zhang, M.-Q., Shao, W., Fan, Y.-Y., Sheng, G.-P., 2019. Mercury/silver resistance genes and their association with antibiotic resistance genes and microbial community in a municipal wastewater treatment plant. Sci. Total Environ. 657, 1014–1022. https://doi.org/10.1016/j.scitotenv.2018.12.088.

Zaatout, N., Bouras, S., Slimani, N., 2021. Prevalence of extended-spectrum β-lactamase (ESBL)- producing Enterobacteriaceae in wastewater: a systematic review and meta-analysis. J. Water Health 19 (5), 705. https://doi.org/10.2166/wh.2021.112.

Zahra, R., Javeed, S., Malala, B., Babenko, D., Toleman, M.A., 2018. Analysis of Escherichia coli STs and resistance mechanisms in sewage from Islamabad, Pakistan indicates a difference in E. Coli carriage types between South Asia and Europe. J. Antimicrob. Chemother. 73, 1781–1785. https://doi.org/10.1093/jac/dky109.

Zankari, E., Allesøe, R., Joensen, K.G., Cavaco, L.M., Lund, O., Aarestrup, F.M., 2017. PointFinder: a novel web tool for WGS-based detection of antimicrobial resistance associated with chromosomal point mutations in bacterial pathogens. J. Antimicrob. Chemother. 72, 2764–2768. https://doi.org/10.1093/jac/dkx217.

Zhang, P., Yang, X., 2022. Genetic characteristics of the transmissible locus of stress tolerance (tLST) and tLST harboring Escherichia coli as revealed by large-scale genomic analysis. Appl. Environ. Microbiol. 88 https://doi.org/10.1128/aem.02185-21 e02185-21.

Zhi, S., Banting, G., Li, Q., Edge, T.A., Topp, E., Sokurenko, M., Scott, C., Braithwaite, S., Ruecker, N.J., Yasui, Y., McAllister, T., Chui, L., Neumann, N.F., 2016. Evidence of naturalized stress-tolerant strains of Escherichia coli in municipal wastewater treatment plants. Appl. Environ. Microbiol. 82, 5505–5518. https://doi.org/10.1128/AEM.00143-16.

Zhi, S., Banting, G., Stothard, P., Ashbolt, N.J., Checkley, S., Meyer, K., Otto, S., Neumann, N.F., 2019. Evidence for the evolution, clonal expansion and global dissemination of water treatment-resistant naturalized strains of Escherichia coli in

wastewater. Water Res. 156, 208–222. https://doi.org/10.1016/j.watres.2019.03.024.

Zhi, S., Stothard, P., Banting, G., Scott, C., Huntley, K., Ryu, K., Otto, S., Ashbolt, N., Checkley, S., Dong, T., Ruecker, N.J., Neumann, N.F., 2020. Characterization of water treatment-resistant and multidrug-resistant urinary pathogenic Escherichia coli in treated wastewater. Water Res. 182, 115827 https://doi.org/10.1016/j.watres.2020.115827.

Zhong, Y.-M., Liu, W.-E., Liang, X.-H., Li, Y.-M., Jian, Z.-J., Hawkey, P.M., 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.