Genomic epidemiology of ESBL-producing *Escherichia coli* from humans and an Aotearoa New Zealand river

1.1 Author names

Holly A Gray (https://orcid.org/0000-0001-5906-2941)[1], Patrick J Biggs (https://orcid.org/0000-0002-0285-4101)[1,2,3], Anne C Midwinter (https://orcid.org/0000-0001-5731-3233)[1], Lynn E Rogers (https://orchid.org/0000-0001-7585-7795)[1], Ahmed Fayaz (https://orcid.org/0000-0001-8879-4535)[1], Rukhshana N Akhter[1], Sara A Burgess (https://orcid.org/0000-0003-1449-7321)[1]#

1.2 Affiliation(s)

[1m]EpiLab, School of Veterinary Science, Massey University, Palmerston North, Aotearoa New Zealand

[2]School of Food Technology and Natural Sciences, Massey University, Palmerston North, Aotearoa New Zealand

[3]New Zealand Food Safety Science and Research Centre, Massey University, Palmerston North, Aotearoa New Zealand

1.3 Corresponding author and email address

Corresponding author: Dr Sara A Burgess, S.Burgess1@massey.ac.nz

1.4 Keywords

ESBL, *E. coli,* antimicrobial resistance, urinary tract infection, freshwater

1.5 Repositories:

All genome data for this study have been deposited in GenBank under accessions PRJNA1032159 and PRJNA715472.

Abstract

In Aotearoa New Zealand, urinary tract infections in humans are commonly caused by extended-spectrum beta-lactamase (ESBL)-producing *Escherichia coli*. This group of antimicrobial resistant bacteria are often multidrug resistant. However, there is limited information on ESBL-producing *E. coli* found in the environment and their link with human clinical isolates. In this study, we examined the genetic relationship of environmental and human clinical ESBL-producing *E. coli* and isolates collected in parallel within the same area over 14 months. Environmental samples were collected from treated effluent, stormwater and multiple locations along an Aotearoa New Zealand river. Treated effluent, stormwater and river water sourced downstream of the treated outflow point were the main sources of ESBL-producing *E. coli* (7/14 samples, 50.0%; 3/6 samples, 50%; and 15/28 samples, 54% respectively). Whole genome sequence comparison was carried out on 307 human clinical and 45 environmental ESBL-producing *E. coli* isolates. Sequence type 131 was dominant for both clinical (147/307, 47.9%) and environmental isolates (11/45, 24.4%). The most prevalent ESBL genes were both $bla_{\text{CTX-M-27}}$ and $bla_{\text{CTX-M-15}}$ for the clinical isolates (134/307, 43.6%) and $bla_{\text{CTX-M-15}}$ for the

environmental isolates (28/45, 62.2%). A core single nucleotide polymorphism analysis of these isolates suggested that some strains were shared between humans and the local river. These results highlight the importance of understanding different transmission pathways for the spread of ESBL-producing *E. coli*.

2. Impact statement

Extended spectrum beta lactamase (ESBL)-producing *E. coli* frequently cause urinary tract infections that exhibit multidrug resistance. Surveillance studies have identified the predominant strains and resistance genes associated with urinary tract infections. However, there is limited information on the extent of spread beyond the patient. We describe the genetic relatedness of ESBL-producing environmental and clinical *E. coli* isolated during the same temporal-spatial period in Aotearoa New Zealand. Comparative genomic analyses of these bacteria provide evidence of clonal spread between humans and the environment, highlighting the need to integrate environmental surveillance into antimicrobial resistance monitoring.

3. Data summary

All Illumina sequence reads for this study have been deposited in GenBank under BioProject PRJNA1032159, except for strain SB0283h1, whose data can be found under BioProject PRJNA715472. The sequence read accessions for each genome are provided in the supplementary material.

The code used for the genomic and statistical analyses is available from the GitHub repository https://github.com/sburgess1/Manawat-_ESBL.

The authors confirm all supporting data and protocols have been provided within the article or through supplementary data files.

4. Introduction

Escherichia coli are gram-negative bacteria that form a natural part of the mammalian intestinal tract microbiota. They can be opportunistic pathogens and can cause a range of infections in humans such as pneumonia, bacteremia, meningitis, and urinary tract infections [1-3]. The spread of antimicrobial resistance in pathogenic strains of *E. coli* is of importance because infections caused by these strains are often harder to treat, resulting in increased severity and duration of infection [4, 5]. One important group of antimicrobial resistant *E. coli* is the extended-spectrum beta-lactamase (ESBL)-producing *E. coli*. In Aotearoa New Zealand (NZ), *E. coli* was associated with 84.6 % of ESBL-producing Enterobacterales related infections in 2019 [6].

The ESBL enzymes confer resistance to an extended range of beta-lactam antibiotics, including the first, third and some fourth generation cephalosporins but not the cephamycins or carbapenems [7, 8]. CTX-M is the predominant enzyme type associated with human clinical ESBL-producing *E. coli* [9-11]. There are more than 250 $bla_{\text{CTX-M}}$ gene variants, with $bla_{\text{CTX-M-15}}$ being the most common amongst human clinical *E. coli* [12-15]. The genes encoding ESBLs are often associated with other

antimicrobial resistance genes (ARGs), resulting in a multidrug resistant (MDR) strain [10, 11, 16]. MDR pathogenic strains of *Enterobacteriaceae* are of particular concern to the health sector resulting in ESBL-producing *Enterobacteriaceae* being identified as "critical" on the World Health Organisation's "Priority Pathogens List" [17].

The main route for community dissemination of antimicrobial-resistant bacteria involves person-to-person transmission [18, 19]. However, there are alternative transmission pathways, such as contact with companion and livestock animals, consumption of food products, or indirect exposure through contaminated recreational water [20-23]. The natural environment and particularly recreational water have been identified as important vectors in the transmission of antimicrobial resistant bacteria [24, 25]. Recreational water use has rarely been found to be a risk factor for ESBL infection and the carriage of ESBL-producing *E. coli* in recreational water users has not been found to be higher compared to non-recreational water users [26-29]. However, freshwater has been identified as a source of ESBL-producing *E. coli* [30, 31]. There are few studies that have compared ESBL-producing *E. coli* from human clinical samples and freshwater samples [32-34]. The aim of this study was to compare the genetic relatedness of clinical and environmental ESBL-producing *E. coli*. Whole genome sequence analysis was carried out for isolates collected over the same time period and geographical region.

5. Methods

5.1 Sample collection and processing

Environmental samples were collected from the Manawatū River (sites A, B, D and F), stormwater (site C), and treated effluent (site E) over 14 months: August 2019 to March 2020 and July 2020 to January 2021 (excluding October 2019). Two water samples and one sediment sample were collected from locations A (upstream of the city of Palmerston North), B, D, and F (downstream) along the Manawatū River (Figure 1). A stormwater sample was collected from the Centennial Drive site (site C) when the stormwater flowed from the drain.

Water samples prepared as previously described [35] and sediment (1 g) were enriched in 10 ml of buffered peptone water (BPW; BD Difco™, Becton Dickinson) or *E. coli* broth (ECB; Oxoid Ltd) and incubated overnight at 35°C. Additionally, 1 ml of treated effluent was diluted in 9 ml of BPW and/or ECB and incubated overnight at 35°C. Two millilitres of each enrichment was centrifuged for 5 min at 6000 g and the pellet was resuspended in 1ml of in-house prepared nutrient broth no. 2 (Oxoid Ltd) containing 15% (v/v) glycerol then stored at -80°C. Each sample was allocated a unique identifying number with the prefix "SB" followed by the allocated sample number. The enrichments were plated onto two selective MacConkey agar plates (containing 1 µg/ml of cefotaxime sodium or ceftazidime) and CHROMagar™ ESBL (Fort Richard Laboratories) as previously described [36]. Additionally, 1 ml of treated effluent was plated onto CHROMagar™ ESBL. After incubation at 35°C for 16 to 24 hours, presumptive *E. coli* colonies (up to two colonies from each plate) were purified on Columbia horse blood agar (Fort Richard Laboratories) and identified using matrix-assisted laser desorption ionization-time of flight (MALDI-TOF Biotyper®, Bruker Daltonics) [36, 37]. Presumptive ESBL-producing *Enterobacteriaceae* clinical strains, which originated from urine, blood, or wound swab cultures were isolated by MedLab Central Laboratory (Palmerston North, Aotearoa New Zealand),

from August 2019 to March 2020 and June 2020 to January 2021. These isolates were also purified on Columbia horse blood agar, identified using MALDI-TOF mass spectrometry, and allocated a unique identifying number with the prefix "EH".

5.2 Antimicrobial susceptibility testing

Antimicrobial susceptibility tests and confirmation of an ESBL-producing phenotype were carried out using Kirby-Bauer disc diffusion assays following Clinical and Laboratory Standards Institute guidelines [38]. Antimicrobial susceptibility tests of the environmental *E. coli* isolates were carried out using a panel of ten antibiotics (MAST Group, Table S1). ESBL confirmation was carried out on the clinical *E. coli* as well as those environmental *E. coli* that were non-susceptible to cefotaxime or ceftazidime using the double-disc comparison assay (D62C cefotaxime and D64C ceftazidime ESBL disc tests; Mast Group).

5.3 DNA Extractions and whole genome sequencing

ESBL-producing *E. coli* isolates were grown in Luria-Bertani broth (made with deionised water, NaCl (Scharlau), yeast extract (Becton Dickinson Difco™) and tryptone (Becton Dickinson Difco™,)) overnight at 35°C and genomic DNA was extracted using the Wizard® Genomic DNA purification kit (Promega) as previously described [35]. Libraries were generated using the Nextera XT DNA library preparation kit (Illumina Inc) and submitted to the Massey Genome Service (Massey University) for quality checks and sequencing. Whole genome sequencing was performed by Novogene (Singapore) using Illumina HiSeq™ X sequencing (2 x 150 bp paired-end reads). FastQC (v.0.11.9) was used for the quality assessment of the raw reads [39].

5.4 Genome assemblies and analyses

The raw reads were processed using Nullarbor 2 (v.2.0.20191013), using the default parameters and strain *E. coli* EC968 was used as the reference (GenBank accession HG941718.1) [40, 41]. Using this pipeline, the adapters were removed using Trimmomatic (v.0.39), the reads were assembled using SKESA (v.2.4.0) and the assemblies were annotated using Prokka (v. 1.14.6) [42-44]. Quast (v.5.0.2) was used for genome evaluations [45]. Additionally, multi-locus sequence-typing (MLST) was carried out using mlst (v.2.19.0), and resistance and virulence genes were identified using ABRicate (v.1.0.1) with the NCBI database (v.2021-03-27) and VirulenceFinder (v.2021- 03-27) databases respectively. The presence of two or more genes as described by [46] (listed in Table S2) was used to determine whether isolates could be classified as extra-intestinal pathogenic *E. coli* (ExPEC) [46]. PointFinder (v.3.0) was used to detect point mutations associated with antimicrobial resistance in the sequences of the ST131 isolates [47]. The *fimH* type was determined using FimTyper (v.1.0) with default parameters for the isolates EH0318a, EH0378a, EH0389a, and EH0391a [48].

ChewBBACA (v.3.1.2) was used to generate a core-genome MLST (cgMLST) allele profile using a 95% threshold [49], and a distance matrix was generated using cgMLST-dists (v.0.4.0, https://github.com/tseemann/cgmlst-dists) [49]. A neighbour-joining tree was generated using the R package ape (v.5.7-1) [50]. A core SNP analysis was carried out using Snippy-multi (v.4.6.0), employing the internal references EH0395a (ST131), EH0143a (ST69), and EH0294a (ST1193) [51]. Gubbins (v.2.3.1) was applied to all Snippy analyses, with the FastTree (v2.1.11) general time-reversible (GTR)

substitution model being used for phylogenetic tree construction [52]. Default parameters for both Gubbins and FastTree were used. Phylogenetic trees were uploaded to the Interactive Tree Of Life (iTOL, v.6.5.7) for annotation and visualisation [50]. To classify the ST131 isolates into clades, sequence reads of known ST131 clades A, B and C were downloaded from the European Nucleotide Archive: strains MER-56 (SRR5936479), MER-53 (SRR5936492) and MER-25 (SRR5936501) respectively [14].

5.5 Plasmidome analysis

Presumptive plasmid contigs were identified using RFPlasmid (v. 0.0.18) [53] and extracted using a custom Python script (v.3.8.1). The plasmid contigs were annotated using Prokka (v. 1.14.6) and the GFF Prokka-generated files were used as the input for a pangenome analysis using Panaroo (v. 1.3.0), which was run in sensitive mode, using the aligner MAFFT, and a core threshold of 0.95 [54, 55]. The gene presence-absence output from Panaroo was uploaded into Panini (https://panini.cgps.group/ Accessed 27 October 2023) and the resulting csv and dot files were uploaded to Microreact (https://microreact.org/ Accessed 27 October 2023) [56, 57].

5.6 Graphical data display

Phenotypic results were stored in a Microsoft Access database (Microsoft 365, v16.0.1) and analysed using the software R (v.4.1.2, R computing group). The code used is available from the GitHub repository: https://github.com/sburgess1/Manawat-_ESBL. R packages used included ggplot2 (v.3.3.5), lubridate (v.1.8.0), ComplexUpset (v.1.3.1), ggalluvial (v.0.12.3), and easyalluvial (v.0.3.0).

6. Results

6.1 The detection of ESBL-producing and multidrug resistant *E. coli* in freshwater, stormwater, and treated sewage

To evaluate the prevalence and persistence of antimicrobial resistant *E. coli* in a waterway, both water and sediment samples were collected from four locations along the Manawatū River over a 14-month period. ESBL-producing *E. coli* were isolated throughout every season, during periods of both low and high rainfall (Figure S1, Table S3). MDR resistance (resistance to three or more classes) and ESBL-producing *E. coli* were detected from three of the four river sample sites (Table 1), with sample site F (downstream of the treated sewage outlet) having a significantly higher proportion of both ESBL-producing (15/28, 53.6%, 95% CI: 29.8 – 72.2%, p=1.863e-09) and multidrug resistant *E. coli* (11/28, 39.3%, 95% CI: 15.9 - 57.6%, p = 2.078e-06) than the other three sites. To evaluate incoming sources of antimicrobial resistant *E. coli*, samples were also collected from a stormwater drain (site C) and the treated sewage outlet (site E). A similar proportion ($p > 0.05$) of ESBL-producing *E. coli* and MDR *E. coli* was observed for both sites when compared to sample site F.

The environmental *E. coli* isolated from the six sample sites were screened for resistance to ten antibiotics (Figure 2). The most observed resistance phenotypes were to cefotaxime, ceftazidime and cefoxitin as well as cefotaxime and ceftazidime. Resistance to trimethoprim-sulfamethoxazole, which is commonly used for the treatment of UTIs [58], was observed in combination with resistance

to one or more other antimicrobials. A multidrug resistance phenotype was observed in 35/82 (42.7%) isolates.

6.2 Genetic diversity of ESBL-producing environmental *E. coli*

We used whole genome sequencing to investigate the genetic diversity of 45 of the environmental ESBL-producing *E. coli* (Figure 3, Table S4). Nineteen different sequence types (STs) were identified, with ST131 being the predominant ST (11/45, 24.4%) followed by ST1722 (6/45, 13.3%), ST10 (3/45, 6.7%) and ST7476 (3/45, 6.7%). The predominant ESBL coding gene was $bla_{\text{CTX-M-15}}$ (28/45, 62.2%, 95% CI: 48.0 – 76.4%) followed by $bla_{\text{CTX-M-27}}$ (8/45, 17.8%, 95% CI: 6.6 – 29.0%) and $bla_{\text{CTX-M-14}}$ (7/45, 15.6%, 95% CI: 5.0 – 26.2%). Twenty-one of the isolates had a multidrug resistance genotype. The number of ARGs ranged from 1 to 19, covering ten antibiotic classes: beta-lactams, aminoglycosides, rifamycins, tetracyclines, sulphonamides, trimethoprim, amphenicols, macrolides, phosphonic acids, and lincosamides. However, the resistance genotype was not always concordant with the resistance phenotype. For example, trimethoprim/sulfamethoxazole resistant *E. coli* isolates did not aways have a *dfr* and/or a *sul* gene present.

6.3 The genetic relationship of ESBL-producing clinical and environmental *E. coli*

To determine whether there was community spread of ESBL-producing *E. coli* between humans and waterways a comparative genomic analysis was carried out, encompassing 307 human clinical ESBL-producing *E. coli* and the 45 environmental isolates (Table S4). Over the 14 months of sampling there appeared to be no seasonal trend, with ST131 (158/352, 44.9%) being the dominant ST (Figure S2). A cgMLST analysis demonstrated that the environmental isolates were dispersed throughout the phylogeny (Figure 4). There were seven river-only isolate STs (ST156, ST219, ST442, ST542, ST1324, ST1584, and ST2079), two effluent-only STs (ST540 and ST635), and 28 human-only STs. The predominant STs across all the isolates were ST131 (158/352, 44.9%), ST1193 (26/352, 7.4%), ST69 (25/352, 7.1%), ST38 (23/352, 6.5%), ST648 (21/352, 6.0%) and ST998 (15/352, 4.3%). Three STs, ST38, ST131, and ST648, were shared across the three sample types (clinical, effluent and river). Three STs were shared across the clinical and river isolates (ST1193, ST69 and ST998). The predominant ESBL coding genes from the clinical isolates were $bla_{\text{CTX-M-15}}$ (134/307, 43.6%, 95% CI: 38.1 – 49.1%) and $bla_{\text{CTX-M-27}}$ (134/307, 43.6%, 95% CI: 38.1 – 49.1%) followed by $bla_{\text{CTX-M-14}}$ (31/307, 10.1%, 95% CI: 6.7 – 13.5%). The number of ARGs ranged from 1 to 20, covering resistance to ten antimicrobial classes. For the three main STs the predominant ESBL coding gene was $bla_{\text{CTX-M-27}}$ for ST131, $bla_{\text{CTX-M-15}}$ for ST69, and $bla_{\text{CTX-M-27}}$ for ST1193.

The majority of ST131 isolates fell within clade C (103/158, 65.2%) followed by clade A (51/158, 32.3%), with the 4/158 (2.5%) belonging to clade B being all clinical isolates. The $bla_{\text{CTX-M-27}}$ was the dominant ESBL coding gene for both the clade A (38/51, 74.5%) and clade C (61/103, 59.2%) isolates. All of the clade C ST131 isolates harboured five point mutations: two in *gyrA*, two in *parC* and one in *parE*; whereas the majority of clade A (44/51, 86.3%) isolates carried two mutations, one each in *gyrA* and *parE*. The ST131 isolates harboured genes typical of ExPEC and UPEC (Table S7), which included the key genes *papA* (1/158, 0.6%), *papC* (77/158, 48.7%), *afaC* (21/158, 13.3%), *kpsM* (146/158, 92.4%), *iutA* (1/158, 0.6%), *chuA* (158/158, 100%), *fyuA* (158/158, 100%), and *sat* (132/158, 83.5%).

We explored the clonal spread of both clinical and environmental ESBL-producing *E. coli* within the three dominant STs (ST131, ST69, and ST1193) by performing a core SNP comparison (Figures 5 and 6, Tables S5, S6 and S7). There was evidence of clonal spread between humans and the environment for ST131 and ST69, but not ST1193. For the ST131 isolates there was less than ten SNPs difference between the environmental ST131 isolates SB0337h1a and SB0338h1b and ten of the clinical isolates, as well as between SB0405h1a and three of the clinical isolates. Two of the clinical strains were received within 14 days of isolating SB0337h1a and SB0338h1b. For the ST69 isolates there was ten SNPs difference between SB0432h1a and three clinical isolates. All four of these environmental strains (SB0337h1a, SB0338h1b, SB0405h1a, SB0432h1a) were isolated from the Manawatū River downstream of the effluent outlet. Although there was no evidence of clonal spread of ST1193 between humans and the environment, the SNP analysis did suggest there was human to human transmission. Isolates EH0056a and EH0097a were isolated from samples collected on the 8th of October 2019 and 19th of November 2019 respectively, with one SNP difference between them, both containing the $bla_{CTX\text{-}M\text{-}3}$ gene, and the same antimicrobial resistant gene profile. There were 4-5 SNPs difference between the isolates EH0087a, EH0111a, and EH0350a. However, EH0111a was received 29 days after EH0087a and EH0350a was received another 11 months later.

6.4 Plasmid analysis

The plasmidome genes across all 352 genomes were compared (Figure 7. There was some clustering by ST, plasmid Inc type and $bla_{CTX\text{-}M}$ variant. There were also clusters with multiple STs, but the same Inc type and $bla_{CTX\text{-}M}$ variant. To determine the probable origin of the $bla_{CTX\text{-}M}$ genes, a plasmidome analysis was carried out, in which 144 contigs with $bla_{CTX\text{-}M}$ genes were identified as originating from plasmids, of which 11 carried a $bla_{CTX\text{-}M}$ gene and plasmid replicon on the same contig. Three different Inc types (IncFII, IncI and IncB/O/K/Z) were associated with these 11 contigs carrying $bla_{CTX\text{-}M}$ genes.

7. Figures and tables

255 **Table 1: Sample level prevalence of ESBL-producing, antimicrobial and multidrug resistant *E. coli*.**

256

	A			B			C	D			E	F		
Sample type	Sediment	Water	Total	Sediment	Water	Total	Storm water	Sediment	Water	Total	Treated effluent	Sediment	Water	Total
ESBL	0/12 (0.0%)	1/28 (3.6%)	1/40 (2.5%)	0/13 (0.0%)	1/28 (3.6%)	1/41 (2.4%)	**3/6 (50.0%)**	0/11 (0.0%)	0/22 (0.0%)	0/33 (0.0%)	**7/14 (50.0%)**	1/12 (8.3%)	**15/28 (53.6%)**	16/40 (40.0%)
Antimicrobial resistant [a]	2/12 (16.7%)	2/28 (7.1%)	2/40 (5.0%)	2/13 (15.4%)	2/28 (7.1%)	4/41 (9.8%)	**4/6 (66.7%)**	0/11 (0.0%)	0/22 (0.0%)	0/33 (0.0%)	**7/14 (50.0%)**	2/12 (16.7%)	**17/28 (60.7%)**	19/40 (47.5%)
Multidrug resistant	0/12 (0.0%)	1/28 (3.6%)	1/40 (2.5%)	0/13 (0.0%)	1/28 (3.6%)	1/41 (2.4%)	2/6 (33.3%)	0/11 (0.0%)	0/22 (0.0%)	0/33 (0.0%)	6/14 (42.9%)	1/12 (8.3%)	11/28 (39.3%)	12/40 (30.0%)

257 [a] Resistant to at least one of the ten antibiotics used to screen the samples.

258
259
260

Figure 1: Sampling sites A to F along the Manawatū river, located in Palmerston North Aotearoa New Zealand. River flow is the direction of site A to F. Image produced using CorelDRAW® 2019 (Corel Corporation). Base map obtained from Google Maps (2023), available at: https://www.google.com/maps/@-40.3521071,175.6424104,12.87z?entry=ttu (Accessed: 16 August 2023).

Figure 2. Antimicrobial resistance profiles of ESBL-producing *E. coli* against ten antibiotics from sample sites A, B, C, E, and F along the Manawatū river. No antimicrobial resistant isolates were isolated from sample site D.

Figure 3. Core-genome MLST neighbour-joining tree, of 45 environmental *E. coli* isolated from river water, sediment, storm water and treated effluent. The tree was generated using 2918 shared alleles identified using chewBBACA.

Figure 4. Core-genome MLST neighbour-joining tree of 352 clinical and environmental *E. coli*. The tree was generated using 3,053 shared alleles identified using chewBBACA.

Figure 5. Core SNP phylogeny of clinical and environmental ST131 isolates. Maximum likelihood phylogenetic tree of 159 ST131 *E. coli* produced using 2,370 SNPs with EH0395a used as the reference genome. The tree was constructed with FastTree using a maximum-likelihood GTR model and visualised in iTOL.

Figure 6. Core SNP phylogeny of ST69 and ST1193 clinical and environmental *E. coli* isolates. a) Maximum likelihood phylogenetic tree of 25 ST69 *E. coli* produced using 630 SNPs with EH0143a used as the reference genome. b) Maximum likelihood phylogenetic tree of 26 ST1193 isolates produced using 726 SNPs with EH0294a used as the reference genome. The tree was constructed with FastTree using a maximum-likelihood GTR model and visualised in iTOL.

Figure 7. Plasmidome network of the plasmid contigs from 352 clinical and environmental *E. coli*. Each circle denotes the concatenated plasmid contigs from one isolate and is coloured by a) sequence type, b) Inc type, c) ESBL and AmpC variant and d) source of isolate. ASP: aspirate, BAG: bag urine, BCF: aspirate fluid (generally blood) in blood culture bottle, BCU: blood culture, CSU: catheter urine, WS: wound swab.

8. Discussion

Our study found that ESBL-producing *E. coli* were frequently isolated from an Aotearoa New Zealand waterway that passes through an urban environment. These bacteria were isolated from at least one sampling site during 12 out of the 13 time-points. Three of the four sampling sites were associated with an urban land-use, while one site was associated with agricultural land-use. The site with the highest prevalence of ESBL-producing *E. coli* was located downstream of a treated effluent outlet (site F). This is not surprising as it has previously been shown that wastewater treatment is not completely effective at removing antimicrobial resistant bacteria [59-62], and studies have found a higher prevalence of ESBL-producing *E. coli* and other types of resistant *E. coli* immediately downstream of effluent outlets [32, 63-65]. A higher prevalence of antimicrobial resistant bacteria

and their genes has also been detected during periods associated with higher rainfall [66-68]. One reason for this increase could be sewage overflow, which has been shown to be the dominant contributor to the bacterial community in a river downstream of a wastewater treatment plant after high rainfall events [69]. In our study ESBL-producing *E. coli* were isolated throughout the year during periods of low and high rainfall. These bacteria were also isolated from sites upstream of the wastewater treatment plant. In Aotearoa New Zealand, it has previously been established that there are antimicrobial resistant *E. coli* present in freshwater environments [30, 35, 70, 71]. Studies suggest that that the recreational use of contaminated waters may also be an exposure route for ESBL-producing *E. coli* associated infections in humans [21, 25]. Although, the carriage of ESBL-producing *E. coli* has not been found to be higher in recreational water users, previous studies have identified recreational water use as a risk factor for ESBL-producing *E. coli* associated infections [26, 28].

The most common ESBL coding gene variant among *E. coli* isolated from waterways differs between studies. Our study found that $bla_{\text{CTX-M-15}}$ was the dominant gene variant for the treated effluent, and water, which concurs with studies carried out in Europe [32, 33, 72]. In contrast, a study carried out in Brazil found $bla_{\text{CTX-M-2}}$ was the dominant gene variant for ESBL-producing *E. coli* isolated from a waterway downstream of a wastewater treatment plant, whereas $bla_{\text{CTX-M-8}}$ was the main ESBL coding gene associated with *E. coli* isolated from the wastewater plant [73]. Other $bla_{\text{CTX-M}}$ variants commonly harboured by ESBL-producing *E. coli* from freshwater include $bla_{\text{CTX-M-1}}$, $bla_{\text{CTX-M-14}}$, $bla_{\text{CTX-M-27}}$ [74-78].

The most frequently detected ESBL gene types amongst the clinical isolates were $bla_{\text{CTX-M-27}}$ and $bla_{\text{CTX-M-15}}$, with both having a proportion of 134/307 (43.6%). These findings are similar to a survey recently undertaken in Aotearoa New Zealand, which found 68/158 (43.0%) human clinical ESBL-producing *E. coli* isolates carried the $bla_{\text{CTX-M-15}}$ and 64/158 (40.5%) carried the $bla_{\text{CTX-M-27}}$ gene [6]. A regional survey also found a similar proportion of ESBL-producing *E. coli* carrying $bla_{\text{CTX-M-27}}$ compared with $bla_{\text{CTX-M-15}}$, 18/65 (27.7%) and 14/65 (21.5%) respectively [13].

In concordance with other studies [72], other ARGs conferring resistance to a range of other antibiotic classes including aminoglycosides, trimethoprim, sulphonamides, tetracyclines, fluoroquinolones, phenicols, and phosphonic antibiotics were present across both the clinical and environmental isolates. Similar plasmid types were also shared across both the clinical and environmental isolates, with IncF being the most common plasmid type. IncF plasmids frequently harbour $bla_{\text{CTX-M}}$ genes [79]. The close clustering of isolates by their plasmidome across some STs in our study suggests horizontal gene transfer may have occurred between strains. Long-read sequencing of plasmids would be needed to confirm which plasmid Inc types harboured the ESBL coding genes.

Both treated effluent and water samples from the river contained a diverse range of *E. coli* sequence types, in agreement with previous studies [32, 33, 62]. ST131, which is the main lineage associated with UTIs and blood infections, was the dominant type for the effluent, river samples, and clinical isolates. In Aotearoa NZ, ST131 remains the most prominent ST found in clinical specimens as reported by previous surveys [6, 13]. In our study, the majority of ST131 clinical and environmental isolates fell within clade C. In contrast, a recent wastewater surveillance study in Canada found that

clade A was the dominant ST131 clade with only 8.6% of strains belonging to clade C [80]. In our study, all the clade C isolates had a fluroquinolone resistance genotype. Elevated levels of fluoroquinolone resistance are generally associated with mutations in both *gyrA* and *parC* [2, 81], and are commonly associated with the C2 clade but is reported to be rare in clade A strains, as was also found in our study [82]. However, a recent study found 72.7% of ST131 clade A strains isolated from wastewater in Canada were resistant to ciprofloxacin [80].

The other dominant ST from the water samples was ST1722. This ST has previously been isolated from humans, sewage, livestock, birds, companion animals and waterways (https://enterobase.warwick.ac.uk/, accessed 13 February 2024). Interestingly, in our study ST1722 was only isolated from water. ST38 and ST648 were isolated from humans, sewage, and river samples. ST38 and ST648 are often associated with blood infections and UTIs in humans and have frequently been isolated from waterways and sewage [32, 33, 62, 83-85]. A previous study conducted in Sweden also isolated ST38 *E. coli* from water samples downstream of a treated effluent outlet. Other studies that have examined freshwater for the presence of ESBL-producing *E. coli* have found that ST949 and ST10 dominate [33, 72]. ST949 was not isolated from the river and ST10 was isolated from the sediment once during our study.

In this study we found that ESBL-producing *E. coli* found in an Aotearoa NZ waterway were genetically similar to those isolated from human clinical infections, where the difference in the number of SNPs was less than or equal to 10 SNPs between several human and river isolates. However, our epidemiological data only supported the recent transmission between humans to the local river within one set of isolates, where the dates of isolation of both the human and the river isolates were within 14 days. The close clustering of multiple human clinical isolates also suggests that there was human-to-human transfer or contact with the same source of ESBL-producing *E. coli* in the community. A limitation to our study is that we only sampled from one geographical area and within this area the treated effluent and river collection occurred once a month, whereas our clinical isolates were collected weekly. This may have reduced our ability to detect more ESBL-producing *E. coli* from the treated effluent and river that were genetically similar to the clinical isolates.

Previous studies have also indicated that there is the spread of ESBL-producing *E. coli* from humans to freshwater (or *vice versa*) [30, 32, 72]. Fagerström, Mölling [32] compared ESBL-producing *E. coli* UTI isolates to those sourced from fresh water using a cgMLST approach and found that some isolates had less than ten allele differences, suggesting the sharing of strains between humans and the environment. A study conducted in Germany found that ST949 *E. coli* isolates collected from swimming and bathing sites were closely related to human clinical ST949 isolates from Aotearoa New Zealand and Sweden, although the difference in the number of SNPs or allele changes between isolates was not stated [72]. The most likely source of human-associated ESBL producing *E. coli* is through the disposal of treated effluent into our waterways [86].

In conclusion, this study found that ESBL-producing *E. coli* are present in water, sediment, stormwater, and treated effluent samples collected along the Manawatū River. It was shown that while treated effluent is a source of antimicrobial resistant *E. coli*, these resistant bacteria were also present in the Manawatū River upstream of the treated effluent outflow. There was some evidence for the sharing of genetically related ESBL-producing *E. coli* between clinical and environmental

sources. The study was limited by the number of isolates collected and therefore sequenced. More frequent sampling would provide a clearer picture of the genetic relatedness between environmental and human ESBL-producing *E. coli* isolates. Our findings emphasise the importance of including an environmental component to antimicrobial resistance surveillance.

9. Author statements

9.1 Author contributions

H.G: Methodology, formal analysis, investigation, funding acquisition, writing – original draft. P.J.B: Genomic analysis, investigation, supervision, funding acquisition, writing – review and editing A. M: Study conceptualisation, methodology, investigation, supervision, funding acquisition, writing – review and editing. L.R: Methodology, resources, investigation, writing – review and editing. A.F: Data curation, investigation, writing – review and editing. R.A: Methodology, investigation, writing – review and editing. S.B: Study conceptualisation, methodology, investigation, funding acquisition, supervision, formal analysis, writing – original draft.

9.2 Conflicts of interest

The authors declare that there are no conflicts of interest.

9.3 Funding information

Funding for this research was sourced from the Palmerston North Medical Foundation and the Hawke's Bay Research Medical Foundation by SB, HG, AM and PB, as well as from the School of Veterinary Science, Massey University by HG.

9.4 Ethical approval

This project has been evaluated by peer review and judged to be low risk (ethics notification number: 4000021252). Consequently, it has not been reviewed by one of Massey University's Human Ethics Committees. This study used bacterial strains, which were isolated from human samples. No biological material of human origin was obtained for this study and the origin of samples was anonymised.

9.5 Acknowledgements

We thank Medlab Central for supplying the clinical Enterobacterales isolates and Massey Genome Services for their assistance with the whole genome sequencing. We wish to acknowledge the use of New Zealand eScience Infrastructure (NeSI) high-performance computing facilities, as part of this research.

10. References

1. **Bonten M, Johnson JR, van den Biggelaar AHJ, Georgalis L, Geurtsen J et al.** Epidemiology of *Escherichia coli* bacteremia: A systematic literature review. *Clin Infect Dis*, 2020;72:1211-1219. doi: 10.1093/cid/ciaa210.

2. **Fibke CD, Croxen MA, Geum HM, Glass M, Wong E et al.** Genomic epidemiology of major extraintestinal pathogenic *Escherichia coli* lineages causing urinary tract infections in young women across Canada. *Open Forum Infect Dis*, 2019;6. doi: 10.1093/ofid/ofz431.
3. **La Combe B, Clermont O, Messika J, Eveillard M, Kouatchet A et al.** Pneumonia-specific *Escherichia coli* with distinct phylogenetic and virulence profiles, France, 2012–2014. *Emerg Infect Dis*, 2019;25:710. doi: 10.3201/eid2504.180944.
4. **Pallett A, Hand K**. Complicated urinary tract infections: practical solutions for the treatment of multiresistant Gram-negative bacteria. *J Antimicrob Chemother*, 2010;65:iii25-iii33. doi: 10.1093/jac/dkq298.
5. **Pitout JDD, Laupland KB**. Extended-spectrum β-lactamase-producing *Enterobacteriaceae*: an emerging public-health concern. *Lancet Infect Dis*, 2008;8:159-166. doi: https://doi.org/10.1016/S1473-3099(08)70041-0.
6. **Tiong A, Woodhouse R, White R, Winter D, Dyet K**. 2019 Survey of extended-spectrum β-lactamase-producing *Enterobacterales* in New Zealand. https://www.esr.cri.nz/our-research/ngakete/infectious-disease-intelligence/antimicrobial-resistance-amr/2023.
7. **Iredell J, Brown J, Tagg K**. Antibiotic resistance in *Enterobacteriaceae*: mechanisms and clinical implications. *Br Med J*, 2016;352:h6420. doi: 10.1136/bmj.h6420.
8. **Rubin JE, Pitout JDD**. Extended-spectrum β-lactamase, carbapenemase and AmpC producing *Enterobacteriaceae* in companion animals. *Vet Microbiol*, 2014;170:10-18. doi: https://doi.org/10.1016/j.vetmic.2014.01.017.
9. **Bhasin A, Chander Y, Manocha H**. Molecular characterization of bla_{CTX-M}, bla_{TEM} and bla_{SHV} beta lactamases produced by uropathogens. *J Adv Med Med Res*, 2020;32:97-106. doi: https://doi.org/10.9734/jammr/2020/v32i2330721.
10. **Findlay J, Gould VC, North P, Bowker KE, Williams MO et al.** Characterization of cefotaxime-resistant urinary *Escherichia coli* from primary care in South-West England 2017–18. *J Antimicrob Chemother*, 2019;75:65-71. doi: 10.1093/jac/dkz397.
11. **Toombs-Ruane LJ, Marshall JC, Benschop J, Drinković D, Midwinter AC et al.** Extended-spectrum β-lactamase- and AmpC β-lactamase-producing Enterobacterales associated with urinary tract infections in the New Zealand community: a case-control study. *Int J Infect Dis*, 2023;128:325-334. doi: https://doi.org/10.1016/j.ijid.2022.12.013.
12. **Baba H, Kuroda M, Sekizuka T, Kanamori H**. Highly sensitive detection of antimicrobial resistance genes in hospital wastewater using the multiplex hybrid capture target enrichment. *mSphere*, 2023;8:e00100-00123. doi: 10.1128/msphere.00100-23.
13. **Hapuarachchi IU, Hannaway RF, Roman T, Biswas A, Dyet K et al.** Genetic evaluation of ESBL-producing *Escherichia coli* urinary isolates in Otago, New Zealand. *JAC - Antimicrob Resist*, 2021;3:dlab147. doi: 10.1093/jacamr/dlab147.
14. **Harris PNA, Ben Zakour NL, Roberts LW, Wailan AM, Zowawi HM et al.** Whole genome analysis of cephalosporin-resistant *Escherichia coli* from bloodstream infections in Australia, New Zealand and Singapore: high prevalence of CMY-2 producers and ST131 carrying $bla_{CTX-M-15}$ and $bla_{CTX-M-27}$. *J Antimicrob Chemother*, 2018;73:634-642. doi: 10.1093/jac/dkx466.
15. **Critchley IA, Cotroneo N, Pucci MJ, Mendes R**. The burden of antimicrobial resistance among urinary tract isolates of *Escherichia coli* in the United States in 2017. *PLOS ONE*, 2019;14:e0220265. doi: 10.1371/journal.pone.0220265.
16. **Lindblom A, Kiszakiewicz C, Kristiansson E, Yazdanshenas S, Kamenska N et al.** The impact of the ST131 clone on recurrent ESBL-producing *E. coli* urinary tract infection: a prospective comparative study. *Sci Rep*, 2022;12:10048. doi: 10.1038/s41598-022-14177-y.
17. **World Health Organization**. Global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics including tuberculosis: World Health Organization2017.

18. **Ludden C, Raven KE, Jamrozy D, Gouliouris T, Blane B et al.** One Health genomic surveillance of *Escherichia coli* demonstrates distinct lineages and mobile genetic elements in isolates from humans versus livestock. *mBio*, 2019;10:e02693-02618. doi: 10.1128/mBio.02693-18.

19. **Martischang R, Riccio ME, Abbas M, Stewardson AJ, Kluytmans JAJW et al.** Household carriage and acquisition of extended-spectrum β-lactamase–producing Enterobacteriaceae: A systematic review. *Infect Control Hosp Epidemiol*, 2020;41:286-294. doi: 10.1017/ice.2019.336.

20. **Dahms C, Hübner N-O, Kossow A, Mellmann A, Dittmann K et al.** Occurrence of ESBL-producing *Escherichia coli* in livestock and farm workers in Mecklenburg-Western Pomerania, Germany. *PLOS ONE*, 2015;10:e0143326. doi: 10.1371/journal.pone.0143326.

21. **Leonard AFC, Zhang L, Balfour AJ, Garside R, Gaze WH**. Human recreational exposure to antibiotic resistant bacteria in coastal bathing waters. *Environ Int*, 2015;82:92-100. doi: 10.1016/j.envint.2015.02.013.

22. **Liu CM, Stegger M, Aziz M, Johnson TJ, Waits K et al.** *Escherichia coli* ST131-H22 as a foodborne uropathogen. *mBio*, 2018;9. doi: 10.1128/mbio.00470-18.

23. **Sparham SJ, Kwong JC, Valcanis M, Easton M, Trott DJ et al.** Emergence of multidrug resistance in locally-acquired human infections with *Salmonella* Typhimurium in Australia owing to a new clade harbouring $bla_{CTX\text{-}M\text{-}9}$. *Int J Antimicrob Agents*, 2017;50:101-105. doi: https://doi.org/10.1016/j.ijantimicag.2017.02.014.

24. **Huijbers PMC, Flach C-F, Larsson DGJ**. A conceptual framework for the environmental surveillance of antibiotics and antibiotic resistance. *Environ Int*, 2019;130:104880. doi: https://doi.org/10.1016/j.envint.2019.05.074.

25. **Leonard AFC, Morris D, Schmitt H, Gaze WH**. Natural recreational waters and the risk that exposure to antibiotic resistant bacteria poses to human health. *Curr Opin Microbiol*, 2022;65:40-46. doi: https://doi.org/10.1016/j.mib.2021.10.004.

26. **Farrell ML, Chueiri A, O'Connor L, Duane S, Maguire M et al.** Assessing the impact of recreational water use on carriage of antimicrobial resistant organisms. *Sci Total Environ*, 2023;888:164201. doi: https://doi.org/10.1016/j.scitotenv.2023.164201.

27. **Larramendy S, Deglaire V, Dusollier P, Fournier JP, Caillon J et al.** Risk factors of extended-spectrum beta-lactamases-producing *Escherichia coli* community acquired urinary tract infections: a systematic review. *Infect Drug Resist*, 2020;13:3945-3955. doi: 10.2147/idr.S269033.

28. **Søraas A, Sundsfjord A, Sandven I, Brunborg C, Jenum PA**. Risk factors for community-acquired urinary tract infections caused by ESBL-producing *Enterobacteriaceae* –A case–control study in a low prevalence country. *PLOS ONE*, 2013;8:e69581. doi: 10.1371/journal.pone.0069581.

29. **van den Bunt G, van Pelt W, Hidalgo L, Scharringa J, de Greeff SC et al.** Prevalence, risk factors and genetic characterisation of extended-spectrum beta-lactamase and carbapenemase-producing *Enterobacteriaceae* (ESBL-E and CPE): a community-based cross-sectional study, the Netherlands, 2014 to 2016. *Euro Surveill*, 2019;24:1800594. doi: https://doi.org/10.2807/1560-7917.ES.2019.24.41.1800594.

30. **Burgess SA, Moinet M, Brightwell G, Cookson AL**. Whole genome sequence analysis of ESBL-producing *Escherichia coli* recovered from New Zealand freshwater sites. *Microb Genom*, 2022;8. doi: https://doi.org/10.1099/mgen.0.000893.

31. **Dhanji H, Murphy NM, Akhigbe C, Doumith M, Hope R et al.** Isolation of fluoroquinolone-resistant O25b:H4-ST131 *Escherichia coli* with CTX-M-14 extended-spectrum β-lactamase from UK river water. *J Antimicrob Chemother*, 2010;66:512-516. doi: 10.1093/jac/dkq472.

32. **Fagerström A, Mölling P, Khan FA, Sundqvist M, Jass J et al.** Comparative distribution of extended-spectrum beta-lactamase–producing *Escherichia coli* from urine infections and environmental waters. *PLOS One*, 2019;14:e0224861. doi: 10.1371/journal.pone.0224861.

33. **Jørgensen SB, Søraas AV, Arnesen LS, Leegaard TM, Sundsfjord A et al.** A comparison of extended spectrum β-lactamase producing *Escherichia coli* from clinical, recreational water and

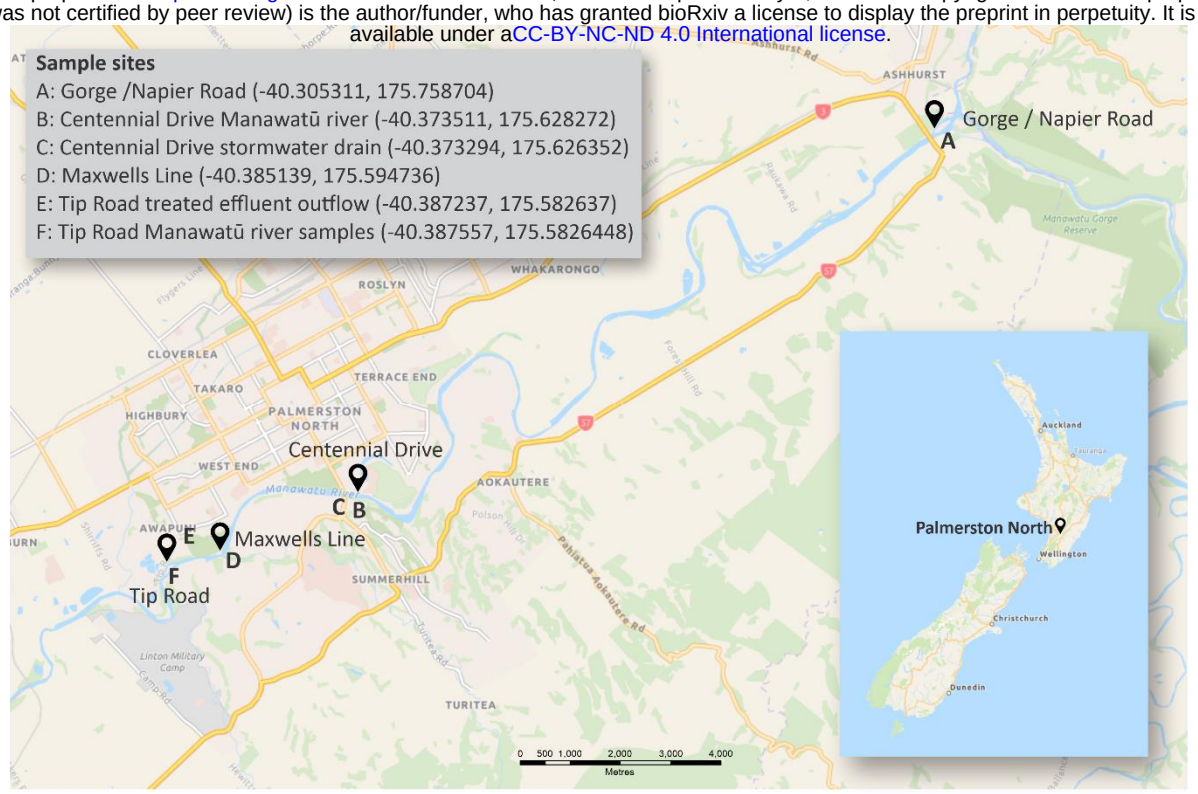
514 wastewater samples associated in time and location. *PLOS One*, 2017;12:e0186576. doi:
515 10.1371/journal.pone.0186576.
516 34. **Adator EH, Walker M, Narvaez-Bravo C, Zaheer R, Goji N et al.** Whole genome sequencing
517 differentiates presumptive extended spectrum beta-lactamase producing *Escherichia coli* along
518 segments of the one health continuum. *Microorganisms*, 2020;8:448. doi:
519 10.3390/microorganisms8030448.
520 35. **Gray HA, Biggs PJ, Midwinter AC, Burgess SA**. Genome sequences for extended-spectrum
521 beta-lactamase-producing *Escherichia coli* strains isolated from different water sources. *Microbiol*
522 *Resour Announc*, 2021;10:e0032821. doi: 10.1128/mra.00328-21.
523 36. **Burgess SA, Aplin J, Biggs PJ, Breckell G, Benschop J et al.** Characterisation of AmpC and
524 extended-spectrum beta-lactamase producing *E. coli* from New Zealand dairy farms. *Int Dairy J*,
525 2021;117:104998. doi: https://doi.org/10.1016/j.idairyj.2021.104998.
526 37. **Schumann P, Maier T**. In: Goodfellow M, Sutcliffe I, Chun J (editors). *Methods in*
527 *Microbiology*: Academic Press; 2014. pp. 275-306.
528 38. **Clinical and Laboratory Standards Institute (CLSI)**. Performance Standards for Antimicrobial
529 Susceptibility Testing. 31st ed. USA: Clinical and Laboratory Standards Institute; 2021.
530 39. **Andrews S**. FastQC. 2010. http://www.bioinformatics.babraham.ac.uk/projects/fastqc.
531 40. **Forde BM, Ben Zakour NL, Stanton-Cook M, Phan MD, Totsika M et al.** The complete
532 genome sequence of *Escherichia coli* EC958: a high quality reference sequence for the globally
533 disseminated multidrug resistant *E. coli* O25b:H4-ST131 clone. *PLOS One*, 2014;9:e104400. doi:
534 10.1371/journal.pone.0104400.
535 41. **Seemann T, Goncalves da Silva A, Bulach D, Schultz M, Kwong J et al.** Nullarbor. 2016.
536 https://github.com/tseemann/nullarbor.
537 42. **Bolger AM, Lohse M, Usadel B**. Trimmomatic: a flexible trimmer for Illumina sequence data.
538 *Bioinformatics*, 2014;30:2114-2120. doi: 10.1093/bioinformatics/btu170.
539 43. **Souvorov A, Agarwala R, Lipman DJ**. SKESA: strategic k-mer extension for scrupulous
540 assemblies. *Genome Biol*, 2018;19:153. doi: 10.1186/s13059-018-1540-z.
541 44. **Seemann T**. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, 2014;30:2068–
542 2069. doi: 10.1093/bioinformatics/btu153.
543 45. **Gurevich A, Saveliev V, Vyahhi N, Tesler G**. QUAST: quality assessment tool for genome
544 assemblies. *Bioinformatics*, 2013;29:1072-1075. doi: 10.1093/bioinformatics/btt086.
545 46. **Johnson JR, Johnston BD, Porter S, Thuras P, Aziz M et al.** Accessory traits and phylogenetic
546 background predict *Escherichia coli* extraintestinal virulence better than does ecological source. *J*
547 *Infect Dis*, 2019;219:121-132. doi: 10.1093/infdis/jiy459.
548 47. **Zankari E, Allesoe R, Joensen KG, Cavaco LM, Lund O et al.** PointFinder: a novel web tool for
549 WGS-based detection of antimicrobial resistance associated with chromosomal point mutations in
550 bacterial pathogens. *J Antimicrob Chemother*, 2017;72:2764-2768. doi: 10.1093/jac/dkx217.
551 48. **Roer L, Tchesnokova V, Allesøe R, Muradova M, Chattopadhyay S et al.** Development of a
552 web tool for *Escherichia coli* subtyping based on *fimH* alleles. *J Clin Microbiol*, 2017;55:2538-2543.
553 doi: 10.1128/JCM.00737-17.
554 49. **Silva M, Machado MP, Silva DN, Rossi M, Moran-Gilad J et al.** chewBBACA: A complete
555 suite for gene-by-gene schema creation and strain identification. *Microb Genom*, 2018;4. doi:
556 10.1099/mgen.0.000166.
557 50. **Letunic I, Bork P**. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree
558 display and annotation. *Nucleic Acids Res*, 2021;49:W293-W296. doi: 10.1093/nar/gkab301.
559 51. **Seemann T**. Snippy: Fast Bacterial Variant Calling from NGS Reads. 2015.
560 https://github.com/tseemann/snippy.
561 52. **Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA et al.** Rapid phylogenetic analysis of
562 large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res*,
563 2014;43:e15-e15. doi: 10.1093/nar/gku1196.

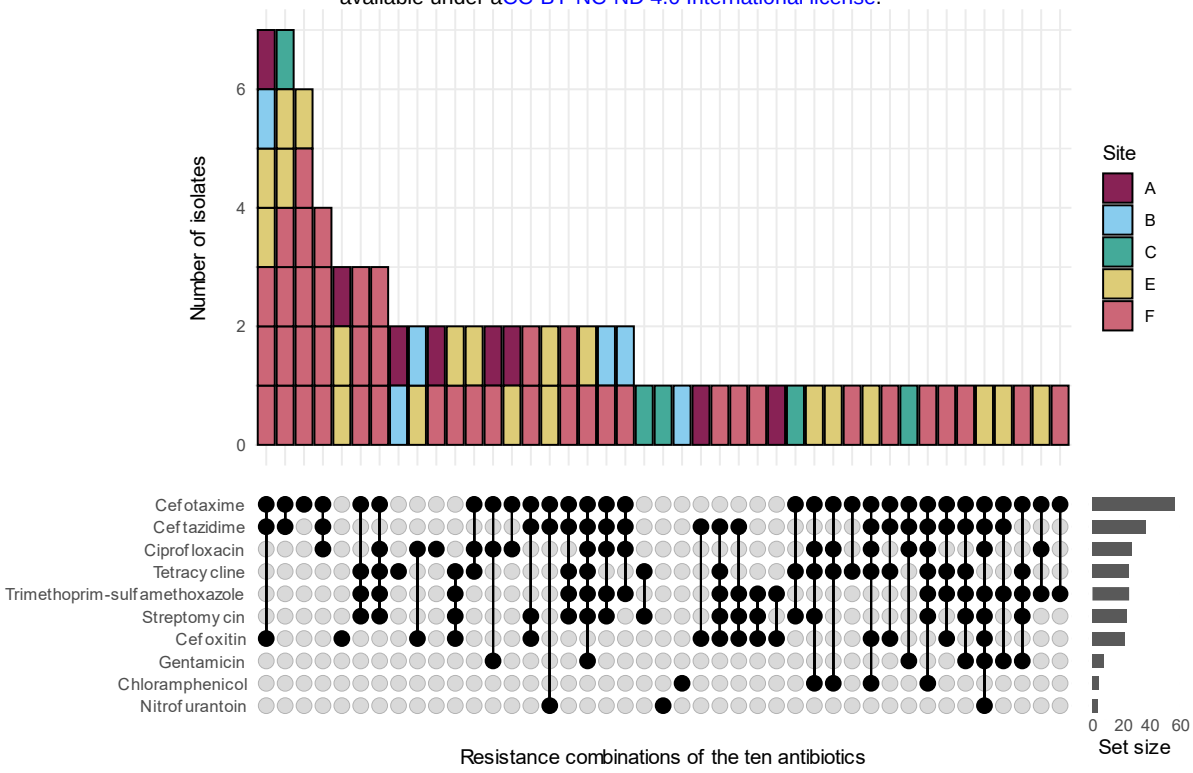
53. **van der Graaf-van Bloois L, Wagenaar JA, Zomer AL**. RFPlasmid: predicting plasmid sequences from short-read assembly data using machine learning. *Microb Genom*, 2021;7. doi: https://doi.org/10.1099/mgen.0.000683.
54. **Tonkin-Hill G, MacAlasdair N, Ruis C, Weimann A, Horesh G et al.** Producing polished prokaryotic pangenomes with the Panaroo pipeline. *Genome Biol*, 2020;21:180. doi: 10.1186/s13059-020-02090-4.
55. **Katoh K, Kuma K, Toh H, Miyata T**. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res*, 2005;33:511-518. doi: 10.1093/nar/gki198.
56. **Argimón S, Abudahab K, Goater RJE, Fedosejev A, Bhai J et al.** Microreact: visualizing and sharing data for genomic epidemiology and phylogeography. *Microb Genom*, 2016;2. doi: https://doi.org/10.1099/mgen.0.000093.
57. **Abudahab K, Prada JM, Yang Z, Bentley SD, Croucher NJ et al.** PANINI: Pangenome Neighbour Identification for Bacterial Populations. *Microb Genom*, 2019;5. doi: https://doi.org/10.1099/mgen.0.000220.
58. **Best Practice Advocacy Centre (BPAC)**. 2017. Antibiotics: choices for common infections. https://bpac.org.nz/antibiotics/bpacnz-antibiotics-guide.pdf [accessed 23 Feb 2022.
59. **McConnell MM, Truelstrup Hansen L, Jamieson RC, Neudorf KD, Yost CK et al.** Removal of antibiotic resistance genes in two tertiary level municipal wastewater treatment plants. *Sci Total Environ*, 2018;643:292-300. doi: https://doi.org/10.1016/j.scitotenv.2018.06.212.
60. **Yuan Q-B, Guo M-T, Yang J**. Fate of antibiotic resistant bacteria and genes during wastewater chlorination: Implication for antibiotic resistance control. *PLOS ONE*, 2015;10:e0119403. doi: 10.1371/journal.pone.0119403.
61. **Fadare FT, Okoh AI**. The abundance of genes encoding ESBL, pAmpC and non-β-lactam resistance in multidrug-resistant *Enterobacteriaceae* recovered from wastewater effluents. *Front Environ Sci*, 2021;9. doi: 10.3389/fenvs.2021.711950.
62. **Raven KE, Ludden C, Gouliouris T, Blane B, Naydenova P et al.** Genomic surveillance of *Escherichia coli* in municipal wastewater treatment plants as an indicator of clinically relevant pathogens and their resistance genes. *Microb Genom*, 2019;5. doi: https://doi.org/10.1099/mgen.0.000267.
63. **Banu RA, Alvarez JM, Reid AJ, Enbiale W, Labi A-K et al.** Extended spectrum beta-lactamase *Escherichia coli* in river waters collected from two cities in Ghana, 2018–2020. *Trop Med Infect*, 2021;6:105. doi.
64. **Mukherjee M, Laird E, Gentry TJ, Brooks JP, Karthikeyan R**. Increased antimicrobial and multidrug resistance downstream of wastewater treatment plants in an urban watershed. *Front Microbiol*, 2021;12. doi: 10.3389/fmicb.2021.657353.
65. **Delgado-Blas JF, Ovejero CM, David S, Montero N, Calero-Caceres W et al.** Population genomics and antimicrobial resistance dynamics of *Escherichia coli* in wastewater and river environments. *Commun Biol*, 2021;4:457. doi: 10.1038/s42003-021-01949-x.
66. **Di Cesare A, Eckert EM, Rogora M, Corno G**. Rainfall increases the abundance of antibiotic resistance genes within a riverine microbial community. *Environ Pollut*, 2017;226:473-478. doi: https://doi.org/10.1016/j.envpol.2017.04.036.
67. **Díaz-Gavidia C, Barría C, Rivas L, García P, Alvarez FP et al.** Isolation of ciprofloxacin and ceftazidime-resistant Enterobacterales from vegetables and river water is strongly associated with the season and the sample type. *Front Microbiol*, 2021;12:604567. doi: 10.3389/fmicb.2021.604567.
68. **Herrig I, Fleischmann S, Regnery J, Wesp J, Reifferscheid G et al.** Prevalence and seasonal dynamics of bla_{CTX-M} antibiotic resistance genes and fecal indicator organisms in the lower Lahn River, Germany. *PLOS ONE*, 2020;15:e0232289. doi: 10.1371/journal.pone.0232289.
69. **Zan R, Blackburn A, Plaimart J, Acharya K, Walsh C et al.** Environmental DNA clarifies impacts of combined sewer overflows on the bacteriology of an urban river and resulting risks to

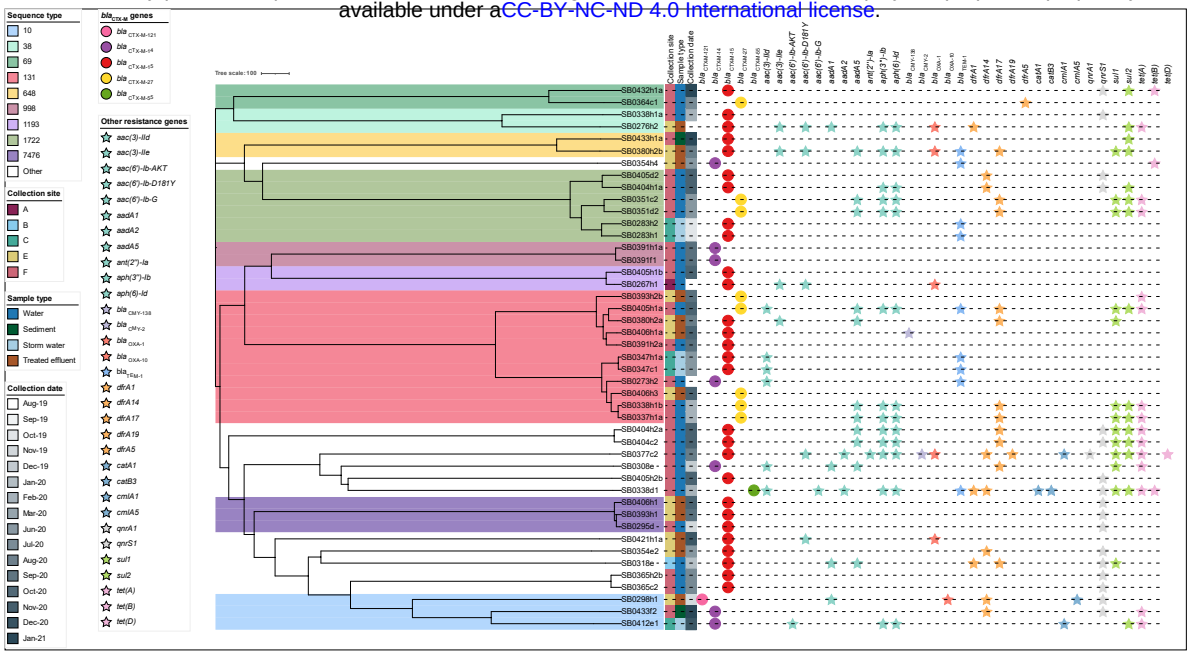
613 public health. *Sci Total Environ*, 2023;889:164282. doi:
614 https://doi.org/10.1016/j.scitotenv.2023.164282.
615 70. **Burgess SA, Francois M, Midwinter AC, Biggs PJ**. Draft genome sequences of seven
616 extended-spectrum β-lactamase-producing *Escherichia coli* strains isolated from New Zealand
617 waterways. *Microbiol Resour Announc*, 2021;10:e01445-01420. doi: 10.1128/MRA.01445-20.
618 71. **van Hamelsveld S, Adewale ME, Kurenbach B, Godsoe W, Harding JS et al.** Prevalence of
619 antibiotic-resistant *Escherichia coli* isolated from urban and agricultural streams in Canterbury, New
620 Zealand. *FEMS Microbiol Lett*, 2019;366. doi: 10.1093/femsle/fnz104.
621 72. **Falgenhauer L, zur Nieden A, Harpel S, Falgenhauer J, Domann E**. Clonal CTX-M-15-
622 producing *Escherichia coli* ST-949 are present in German surface water. *Front Microbiol*, 2021;12.
623 doi: 10.3389/fmicb.2021.617349.
624 73. **Conte D, Palmeiro JK, da Silva Nogueira K, de Lima TMR, Cardoso MA et al.**
625 Characterization of CTX-M enzymes, quinolone resistance determinants, and antimicrobial residues
626 from hospital sewage, wastewater treatment plant, and river water. *Ecotoxicol Environ Saf*,
627 2017;136:62-69. doi: https://doi.org/10.1016/j.ecoenv.2016.10.031.
628 74. **Baron S, Le Devendec L, Lucas P, Larvor E, Jové T et al.** Characterisation of plasmids
629 harbouring extended-spectrum cephalosporin resistance genes in *Escherichia coli* from French rivers.
630 *Vet Microbiol*, 2020;243:108619. doi: https://doi.org/10.1016/j.vetmic.2020.108619.
631 75. **Bollache L, Bardet E, Depret G, Motreuil S, Neuwirth C et al.** Dissemination of CTX-M-
632 producing *Escherichia coli* in freshwater fishes from a French watershed (Burgundy). *Front Microbiol*,
633 2019;9. doi: 10.3389/fmicb.2018.03239.
634 76. **Miyagi K, Hirai I**. A survey of extended-spectrum β-lactamase-producing *Enterobacteriaceae*
635 in environmental water in Okinawa Prefecture of Japan and relationship with indicator organisms.
636 *Environ Sci Pollut Res*, 2019;26:7697-7710. doi: 10.1007/s11356-019-04189-z.
637 77. **Franz E, Veenman C, van Hoek AHAM, de Roda Husman A, Blaak H**. Pathogenic *Escherichia*
638 *coli* producing extended-spectrum β-Lactamases isolated from surface water and wastewater. *Sci*
639 *Rep*, 2015;5:14372. doi: 10.1038/srep14372.
640 78. **Girlich D, Bonnin RA, Naas T**. Occurrence and diversity of CTX-M-producing *Escherichia coli*
641 from the Seine River. *Front Microbiol*, 2020;11. doi: 10.3389/fmicb.2020.603578.
642 79. **de Been M, Lanza VF, de Toro M, Scharringa J, Dohmen W et al.** Dissemination of
643 cephalosporin resistance genes between *Escherichia coli* strains from farm animals and humans by
644 specific plasmid lineages. *PLOS Genetics*, 2014;10:e1004776. doi: 10.1371/journal.pgen.1004776.
645 80. **Finn Thomas J, Scriver L, Lam L, Duong M, Peirano G et al.** A comprehensive account of
646 *Escherichia coli* sequence type 131 in wastewater reveals an abundance of fluoroquinolone-resistant
647 clade A strains. *Appl Environ Microbiol*, 2020;86:e01913-01919. doi: 10.1128/AEM.01913-19.
648 81. **Hopkins KL, Davies RH, Threlfall EJ**. Mechanisms of quinolone resistance in *Escherichia coli*
649 and *Salmonella*: Recent developments. *Int J Antimicrob Agents*, 2005;25:358-373. doi:
650 https://doi.org/10.1016/j.ijantimicag.2005.02.006.
651 82. **Pitout J, DeVinney R**. *Escherichia coli* ST131: a multidrug-resistant clone primed for global
652 domination. *F1000Res*, 2017;6. doi: 10.12688/f1000research.10609.1.
653 83. **Day MJ, Rodríguez I, van Essen-Zandbergen A, Dierikx C, Kadlec K et al.** Diversity of STs,
654 plasmids and ESBL genes among *Escherichia coli* from humans, animals and food in Germany, the
655 Netherlands and the UK. *J Antimicrob Chemother*, 2016;71:1178-1182. doi: 10.1093/jac/dkv485.
656 84. **Paulshus E, Thorell K, Guzman-Otazo J, Joffre E, Colque P et al.** Repeated isolation of
657 extended-spectrum-lactamase-positive *Escherichia coli* sequence types 648 and 131 from
658 community wastewater indicates that sewage systems are important sources of emerging clones of
659 antibiotic-resistant bacteria. *Antimicrob Agents Chemother*, 2019;63. doi: 10.1128/aac.00823-19.
660 85. **Furlan JPR, Savazzi EA, Stehling EG**. Widespread high-risk clones of multidrug-resistant
661 extended-spectrum β-lactamase-producing *Escherichia coli* B2-ST131 and F-ST648 in public aquatic

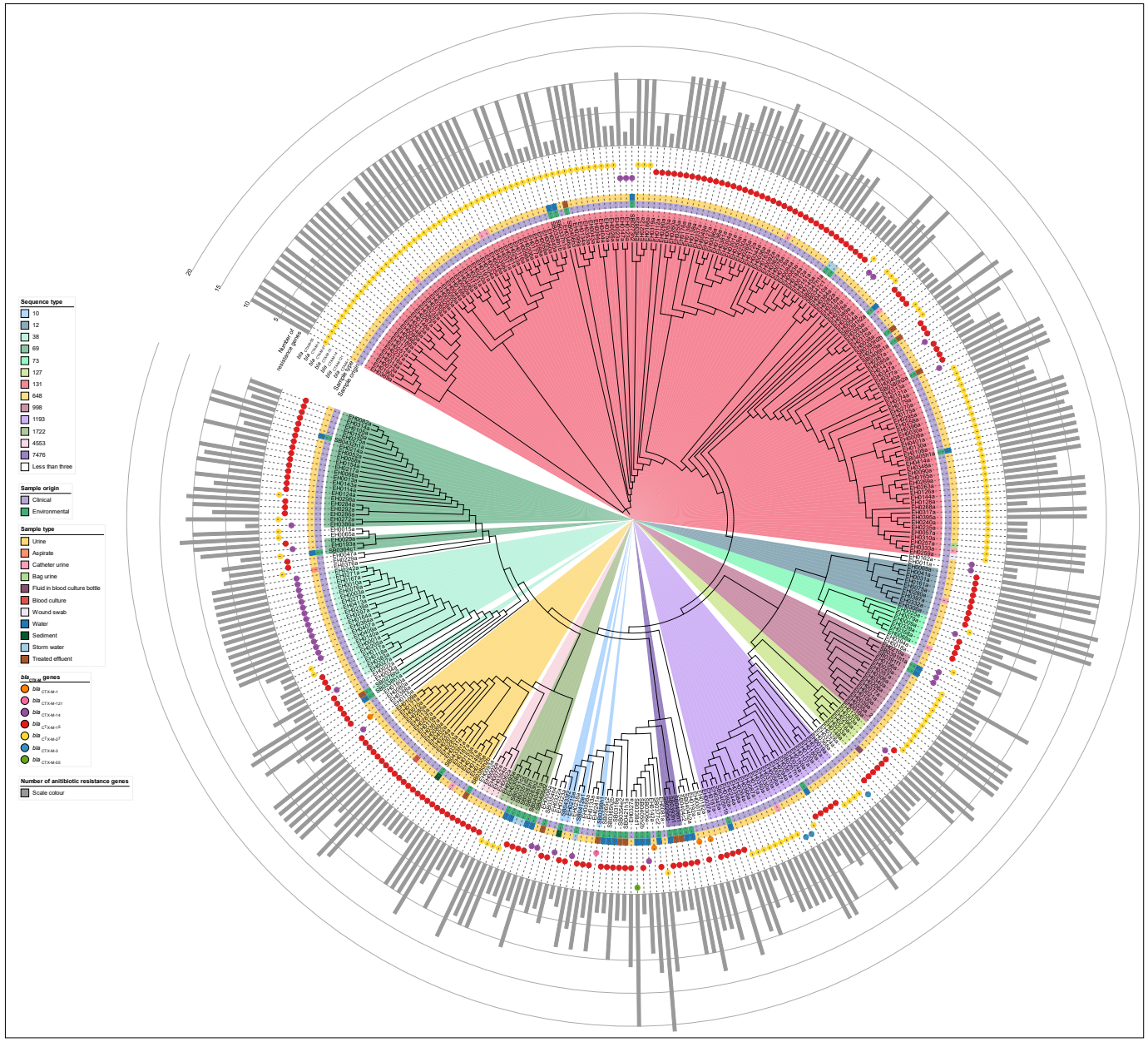
662 environments. *Int J Antimicrob Agents*, 2020;56:106040. doi:
663 https://doi.org/10.1016/j.ijantimicag.2020.106040.
664 86. **Hooban B, Fitzhenry K, O'Connor L, Miliotis G, Joyce A et al.** A longitudinal survey of
665 antibiotic-resistant *Enterobacterales* in the Irish environment, 2019–2020. *Sci Total Environ*,
666 2022;828:154488. doi: https://doi.org/10.1016/j.scitotenv.2022.154488.

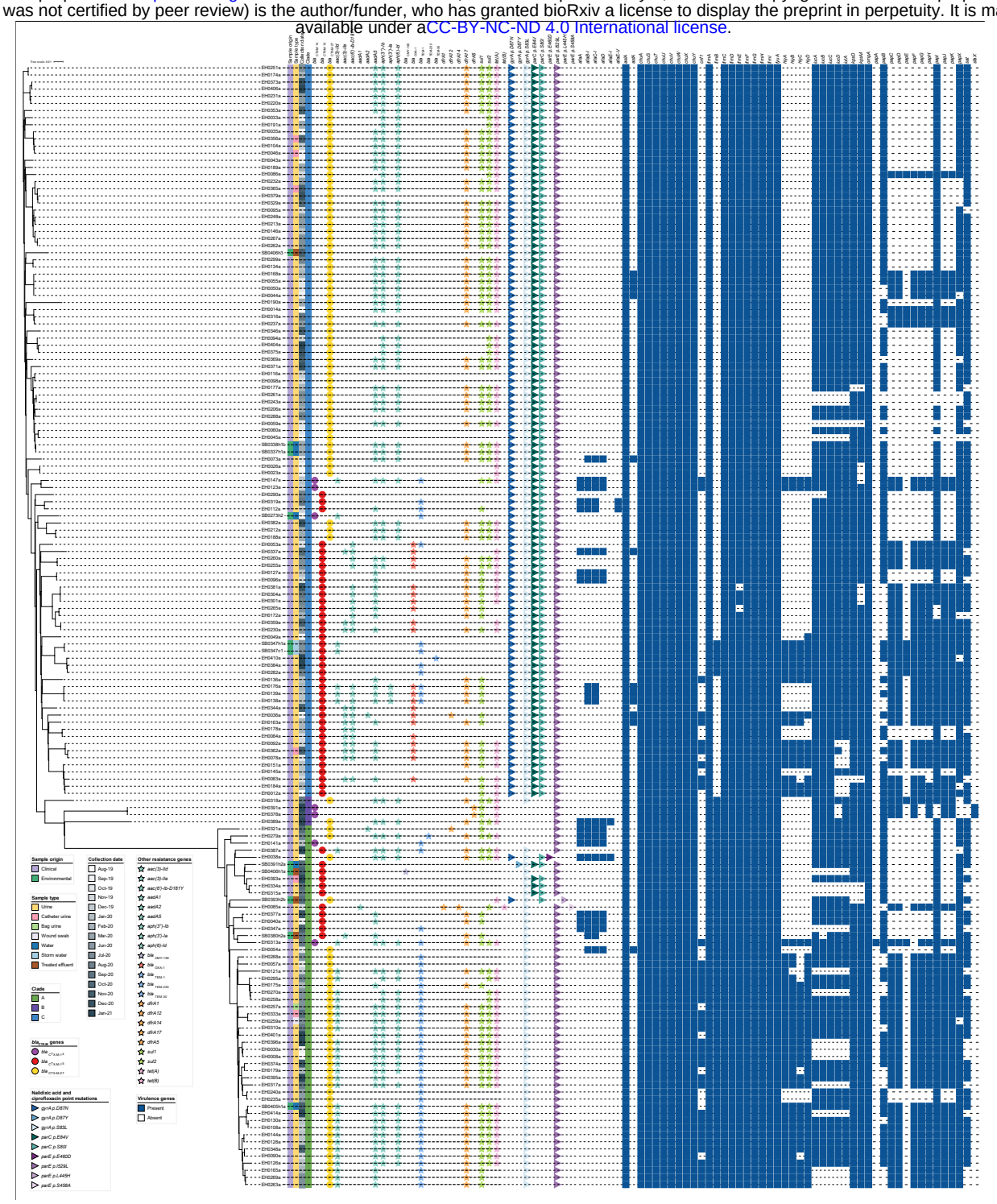
667

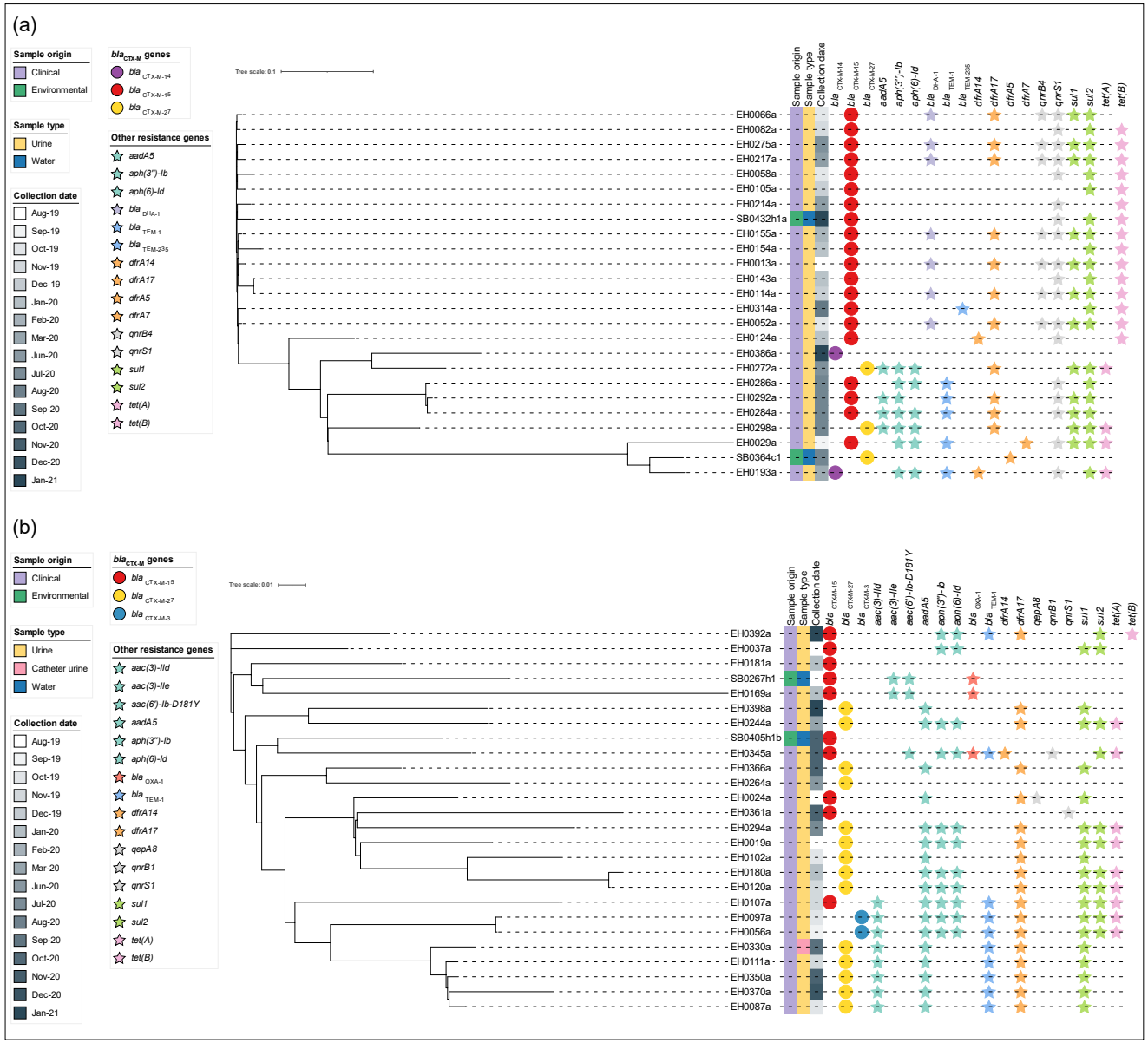


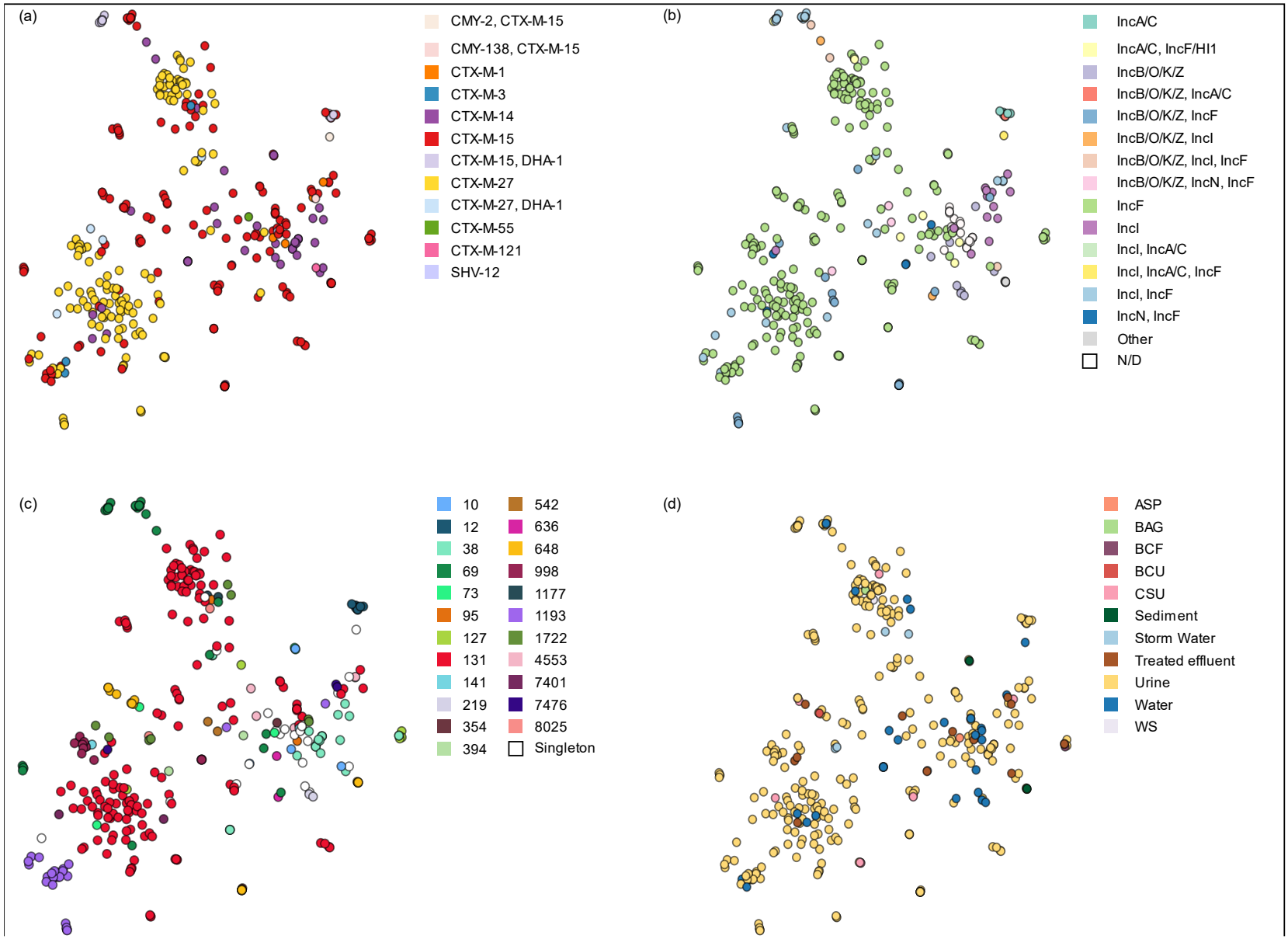





Sequence type
10
12
38
69
73
127
131
648
998
1193
1722
4553
7476
Less than three
Sample origin
Clinical
Environmental
Sample type
Urine
Aspirate
Catheter urine
Bag urine
Fluid in blood culture bottle
Blood culture
Wound swab
Water
Sediment
Storm water
Treated effluent
blaCTX-M genes
blaCTX-M-1
blaCTX-M-121
blaCTX-M-14
blaCTX-M-15
blaCTX-M-27
blaCTX-M-9
blaCTX-M-55
Number of antibiotic resistance genes
Scale colour
Number of resistance genes
blaCTX-M-1
blaCTX-M-121
blaCTX-M-14
blaCTX-M-15
blaCTX-M-27
blaCTX-M-9
blaCTX-M-55
Sample type
Sample origin
5
10
15
20




(a)
Sample origin
Clinical
Environmental
Sample type
Urine
Water
Collection date
Aug-19
Sep-19
Oct-19
Nov-19
Dec-19
Jan-20
Feb-20
Mar-20
Jun-20
Jul-20
Aug-20
Sep-20
Oct-20
Nov-20
Dec-20
Jan-21
blaCTX-M genes
blaCTX-M-14
blaCTX-M-15
blaCTX-M-27
Other resistance genes
aadA5
aph(3")-Ib
aph(6)-Id
blaOXA-1
blaTEM-1
blaTEM-235
dfrA14
dfrA17
dfrA5
dfrA7
qnrB4
qnrS1
sul1
sul2
tet(A)
tet(B)
Tree scale: 0.1
EH0066a
EH0082a
EH0275a
EH0217a
EH0058a
EH0105a
EH0214a
SB0432h1a
EH0155a
EH0154a
EH0013a
EH0143a
EH0114a
EH0314a
EH0052a
EH0124a
EH0386a
EH0272a
EH0286a
EH0292a
EH0284a
EH0298a
EH0029a
SB0364c1
EH0193a
(b)
Sample origin
Clinical
Environmental
Sample type
Urine
Catheter urine
Water
Collection date
Aug-19
Sep-19
Oct-19
Nov-19
Dec-19
Jan-20
Feb-20
Mar-20
Jun-20
Jul-20
Aug-20
Sep-20
Oct-20
Nov-20
Dec-20
Jan-21
blaCTX-M genes
blaCTX-M-15
blaCTX-M-27
blaCTX-M-3
Other resistance genes
aac(3)-IId
aac(3)-IIe
aac(6')-Ib-D181Y
aadA5
aph(3")-Ib
aph(6)-Id
blaOXA-1
blaTEM-1
dfrA14
dfrA17
qepA8
qnrB1
qnrS1
sul1
sul2
tet(A)
tet(B)
Tree scale: 0.01
EH0392a
EH0037a
EH0181a
SB0267h1
EH0169a
EH0398a
EH0244a
SB0405h1b
EH0345a
EH0366a
EH0264a
EH0024a
EH0361a
EH0294a
EH0019a
EH0102a
EH0180a
EH0120a
EH0107a
EH0097a
EH0056a
EH0330a
EH0111a
EH0350a
EH0370a
EH0087a


(a)
CMY-2, CTX-M-15
CMY-138, CTX-M-15
CTX-M-1
CTX-M-3
CTX-M-14
CTX-M-15
CTX-M-15, DHA-1
CTX-M-27
CTX-M-27, DHA-1
CTX-M-55
CTX-M-121
SHV-12
(b)
IncA/C
IncA/C, IncF/HI1
IncB/O/K/Z
IncB/O/K/Z, IncA/C
IncB/O/K/Z, IncF
IncB/O/K/Z, IncI
IncB/O/K/Z, IncI, IncF
IncB/O/K/Z, IncN, IncF
IncF
IncI
IncI, IncA/C
IncI, IncA/C, IncF
IncI, IncF
IncN, IncF
Other
N/D
(c)
10
12
38
69
73
95
127
131
141
219
354
394
542
636
648
998
1177
1193
1722
4553
7401
7476
8025
Singleton
(d)
ASP
BAG
BCF
BCU
CSU
Sediment
Storm Water
Treated effluent
Urine
Water
WS