40328586

Identification and whole genome analysis of cefotaxime resistant *Escherichia coli* recovered from a tertiary treatment wastewater treatment plant discharge point.

LI Black

Thesis submitted in partial fulfilment of the requirements of Edinburgh Napier University, for the award of Masters by Research

February 2024

Supervisors: Dr Fiona M Stainsby and Dr Donald Morrison

40328586

Declaration

I LI Black, hereby certify that this thesis and the work presented in it are the result of my own original research. No part of this work has been submitted for any degree other than Master by Research at Edinburgh Napier University. Where other sources of information have been used, they have been duly acknowledged.

Name: LI Black

Date: 10 February 2024

Abstract

Wastewater treatment plants (WWTPs) act as reservoirs for antibiotic resistant bacteria (ARB) and may facilitate transmission of antibiotic resistance genes (ARGs) between bacterial populations including pathogens that in turn may persist in wider aquatic environments. ESBL-producing *Escherichia coli* (*E. coli*) cause both hospital and community-acquired infections globally and may serve as a proxy of a wider problem. The aim of this research was to investigate the abundance and diversity of ESBL-producing *E. coli* released from a municipal WWTP to the river Almond in West Lothian, Scotland. Further, whole genome sequence (WGS) analysis of ESBL-producing *E. coli* isolates was undertaken to gain insight to the potential contribution of this WWTP to the resistome in the adjacent aquatic environment.

A culture-based study found approximately 2.0 x10^2 CFU ml^{-1} cefotaxime resistant *E. coli* in the effluent outflow from the WWTP. Isolates were confirmed as *E. coli* by *gad*A and *uid*A targeted primers and further characterised to five *E. coli* phylogroups (B1, B2, C, D, and F) and 5 REP-PCR types. Representatives of each REP type were confirmed to carry Group 1 and 9 CTX-M enzymes. All but 1 isolate exhibited multi drug resistance (MDR) when challenged with representatives from different antibiotic classes. Whole genome sequencing (WGS) and analysis was performed on four resistant *E. coli* isolates. ARGs, mobile genetic elements, and genetic lineages were identified using web-based bioinformatic tools. All resistant strains were confirmed to carry CTX-M genes. ARGs conferring resistance to beta-lactams, aminoglycosides, macrolides, fluoroquinolones, tetracyclines, folate pathway agonists, phenicols, and disinfectants were also identified amongst the resistant strains. ESBL resistance was mobilised, either by IS6 family composite transposons maintained in IncFII or Inc1I plasmids. Other MDR regions were mobilised by Int1 integrons. Importantly, 3 of the 4 isolates belong to globally recognised lineages associated with bacteraemia in humans, including sequence type (ST)131 and ST69. In conclusion, the presence of ESBL-producing, and indeed globally important multi-drug resistant *E.coli* strains in the effluent indicate a potential role of the WWTP in the dissemination of ARB. These findings support the case for active surveillance of ARB and ARGs released from municipal WWTPs.

Acknowledgements

The main debt of gratitude is directed towards the sponsors of this Thesis, Derek and Maureen Moss, whose kindness allowed me to undertake this work.

Further gratitude is directed towards both my supervisors for their timely feedback and encouragement. Their patience with my shifting to a part-time commitment was hugely appreciated.

Finally, a thank you to the lab technicians at Napier who would periodically go out of their way to facilitate me and always complimented the laboratory with an element of cheer.

"And therefore, think him as a serpent's egg / Which hatched, would as his kind grow mischievous"

Table of Contents

List of Figures

List of Tables

Abbreviations

C	Celsius
CTX	Cefotaxime
DNA	Deoxyribonucleic acid
g	gram
h	hour
ml	millilitres
mm	millimetre
NA	Nutrient Agar
PCR	Polymerase chain reaction
sp. (spp.)	species (species—plural)
µl	microlitre
ND	None detected
pmol	Picamol
MDR	Mult- drug resistance
MLST	Multi-locus Sequence Typing
WGS	Whole Genome Sequencing
bla	Beta-lactamase
LGT	Lateral gene transfer
MGE	Mobile genetic element
IS	Insertion sequence
ESBL	Extended Spectrum Beta Lactamase
CGE	Centre for Genomic Epidemiology
WWTP	Wastewater Treatment Plant

1. Introduction

1.1. Antibiotics and antibiotic resistance.

Antimicrobials are ancient bioactive molecules that are ubiquitous throughout the natural environment (Davies and Davies, 2010). They are produced as secondary metabolites by microorganisms as a response to competition and predation (Ageitos *et al*., 2017; Khalil *et al.,* 2019), and facilitate niche exploitation with producers selectively acquiring resistance to their own production (Gillings, 2013). At lower concentrations, these compounds are also involved in bacterial cell-to-cell signalling (Linares *et al.,* 2006).

1.1.1. Historical perspective.

Resistance to antimicrobials is as ancient as the biosynthesis of antimicrobials themselves (Perry *et al.,* 2016) due to both defensive response and as a response to self-toxicity. Some β-lactamases are estimated to be a least 2 billion years old with mobilisation possibly occurring millions of years ago. What is important though is that mechanisms that provide resistance to antibiotics are present in the absence of antibiotic selection, suggesting alternative putative functions associated with these genes (Diene *et al.,* 2020). Darwinian selection has favoured the development of multi-resistance mechanisms against these molecules with pre-existing adaptions being modified rapidly through mutations, recombination and becoming mobile by gene transfer events, such as lateral gene transfer (LGT) (Davies and Davies, 2010; Jansen and Aktipia, 2014).

Although in the modern clinical sense, the first antibiotic discovery is attributed to Fleming in 1929, a plethora of natural phytochemicals, including the moulds that produce penicillin, have found traditional uses as antimicrobials by various civilisations for millennia (Guzman *et al.,* 2012). Adoption of antimicrobials into the clinical mainstream occurred shortly after germ theory was first linked with algae and fungi (the latter once postulated as being responsible for cholera) and then popularised with bacteria through the initial identification of microorganisms as the causative agents of infectious diseases by Louis Pasteur (Wainwright, 2008; Berche, 2012; Smith, 2012). Around this time, microbial antagonism began to be recognised as demonstrated by the inhibitory effects of *Penicillium glaucum* on bacterial growth by T.H Huxley, an observation not investigated again until the serendipitous

meanderings of fungi favoured Fleming in 1928 (Friday, 1974). However, the use antibiotics has come with a substantial cost with bacterial cells developing resistance to antibiotics used widely in clinical settings through mutations and uptake of foreign DNA, an occurrence observed by Fleming almost immediately (Fleming, 1929).

In the present day, antimicrobial resistance is considered a global pandemic. A final report by O'Neill (2016) commissioned by the UK prime minister David Cameron in 2014, suggested AMR to be responsible for 700,000 deaths globally a year, a number projected to reach 10 million by 2050 (O'Neill, 2016). These findings were mirrored by the United Nations declaration on AMR (UN 2016) and the EU AMR action plan of 2017 (EU 2017) and act as a part of the WHO One Health Global Action Plan encouraging collaboration between countries (WHO 2017). Since the report by O'Neill (2016), it has been predicted that AMR was associated with an estimated 4.95 million deaths in 2019 with drug resistant bacteria, namely the ESKAPE priority pathogens, comprising of about 1.27million (26%) of these deaths (Murray *et al*., 2022).

1.1.2. Penicillin and the β-lactams.

Penicillin G, also known as benzylpenicillin, was the first penicillin antibiotic to be produced and together with penicillin V (phenoxymethylpenicillin) targeted streptococcal and staphylococcal infections. However, widespread resistance to these natural penicillins quickly emerged in staphylococci that produced penicillinase (beta-lactamase), an enzyme that hydrolytically cleaves the beta-lactam ring of penicillin (Figure 1a) inactivating its efficacy. Indeed, Fleming noted resistance against penicillin in his initial paper in 1929 and again in 1940 (Abraham and Chain, 1940). This led to the development of new penicillinase resistant penicillins or semi-synthetic penicillins (Figure 1c).

This surge in development subsequently initiated the development and clinical use of the β-lactam antibiotics and two strategies to combat β-lactamase activity emerged. Firstly, the mass screening for β-lactams and synthetic adjustments to β-lactams as a method to evade β-lactamase enzymatic activity. Secondly the development of inhibitors to actively target B-lactamases, with the latter to be used in combination with the former (Papp-Wallace, 2019).

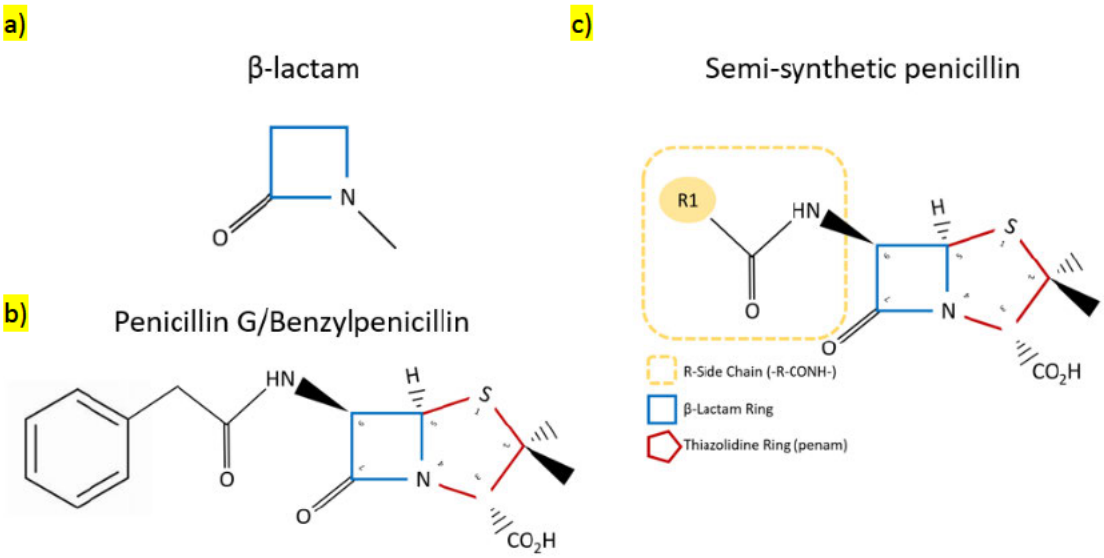


Figure 1.1 a) 4 membered 2-azetidinone ring (β-lactam ring) as the structural core of (b) Benzylpenicillin (6-aminopenicillinic acid). It comprises of 5 membered thiazolidine ring and an acyl-amino side chain attached to the 6-position carbon. c) Synthetic adjustments at position 6, facilitates the addition of synthetic side chains.

While penicillin was the first clinically applied β-lactam, the β-lactam market until the 1970's comprised of penicillins followed by the cephalosporins (initially isolated from *Acremonium chrysogenum*). Even though similarly structured utilising the 4 membered 2-azetidinone ring, (β-lactam ring), the cephalosporins were identified with a fused 6 membered heterocyclic ring as opposed to the 5 membered heterocyclic ring in penicillin (Hamed *et al.,* 2013).

Structural differences define the β-lactams into their family groups (Hamed *et al.*, 2013). All β-lactams comprise of the 4-membered β-lactam ring which may be either fused with 5 membered (thiazolidine) 6 membered (dihydrothiazine) heterocyclic rings. The heteroatom at position 1 may either be a sulphur (penicillins and cephalosporins), an oxygen (clavams), or a carbon atom (carbapenems). The one exception to this are the monocyclic β-lactams, which as the name suggests is solely structured around the β-lactam ring and do not have a fused thiazolidine or dihydrothiazine ring (Hamed *et al.*, 2013; Bush and Bradford, 2016).

1.1.3. β-lactamases.

β-lactamases were commonly regarded as a direct bi-product of the modern antibiotic era. However, exploration in modern times have identified unique gene sequences in bacterial species that have been isolated for millions of years. Further to this is the observation that they are not unique to bacteria as they have been identified in fungi as well. β-lactamase genes are potentially 2 billion years old and are suggested to have become mobilised from chromosomes onto plasmids from as early as 500 million years ago and evolved simultaneously with β-lactamase producing microorganisms as a counter measure to β-lactams production through co-evolution with other β-lactam producing microorganisms as a counter measure to β-lactamase production. This pre-clinical observation of β-lactamase activity is further demonstrated by genomics which consistently find unique β-lactamase type genes without homology to the commonly identified β-lactamases (Bush, 2018).

Despite the ancient origins of β-lactamases, the discovery has no doubt been expediated through their rapid deleterious effect on efficacy of clinically administered β-lactams. Activity of β-lactamases functions through the hydrolysis of the amide bond in β-lactams. This effectively prevents the β-lactam from inhibiting cellular penicillin binding domain function. β-lactamase hydrolytic functionality mainly occurs through one of two mechanisms. Firstly, through an active site serine residue that utilizes an acylation-deacylation mechanism, and secondly, through zinc ions in metallo-β-lactamase active sites that utilises different mechanisms depending upon the enzyme. Recognition of these two different functionalities has subsequently facilitated the broad classification of β-lactamases into four classes, A-D, with Class A, C and D being the serine β-lactamases and Class B being the metallo-β-lactamases, although more recently these four groups can be broken down into 17 functional groups and even further using sequencing data (Drawz and Bonomo, 2010; Hamed *et al*., 2013; Tooke *et al*., 2019).

Serine β-lactamases share an almost identical structure to the serine transpeptidases (the main cross-linking enzymes in most bacteria), sharing the serine-xxaa-xxaa-lysine motif although they lack the serine peptidase activity. However, despite the conserved residues in Penicillin Binding Proteins (PBP's), functional behaviour is not similarly conserved. For example, while both enzymes have a rapid acylation step the deacylation rate of

transpeptidase PBP's mentioned above is slow and conversely it is rapid in β-lactamases (Hamed *et al*., 2013).

The class A β-lactamases are the most widely studied and include the PC1, KPC, TEM, SHV and CTX-M enzymes. The TEM, SHV and CTX-M enzymes have rapidly become notorious in the Enterobacteriaceae. The precipitous expansion in spectrum of activity due to the enzymes swift mutational adaption to new clinical substrates has resulted in an additional nomenclature to Extended Spectrum β-lactamase (ESBL). Not only have these class A -β-lactamases expanded their spectrum of activity but are observed to possess an affinity with Mobile Genetic Elements (MGE's) resulting in their global distribution. A recent addition to the clinical arsenal were the carbapenems which functioned to acetylate most of the serine β-lactamases including the many of the class A β-lactamases. The KPC enzymes can effectively hydrolyse carbapenems owing to structural variability due to an expanded active site which increases hydrophobicity effectively prolonging the hydrogen-bonding interaction long enough for deacetylation to occur. Again, these KPC enzymes (although mostly restricted to *Klebsiella pneumoniae*) are globally renowned as the most prevalent Class A carbapenemases (Drawz and Bonomo, 2010; Hamed *et al*., 2013; Bush and Bradford, 2016; Tooke *et al*., 2019).

The Class D enzymes have the OXA nomenclature and are all Oxacillinases. Originally activity was restricted to penicillins, but more recently the spectrum of activity has expanded to include both cephalosporins and carbapenems and currently have a broad host range, disseminating into both Gram positive and negative bacteria (Tooke *et al*., 2019). OXA enzyme activity is identified as lysine carboxylation in the active site and similarly to the KPC enzymes, the OXA enzymes have an increased hydrophobicity in this active site, due to increased non-polar surface area in comparison to the other Class A enzymes. Again, similarly to KPC activity against carbapenems is demonstrated to be due to steric constraints upon the carbapenem acylenzyme orientation. It is now recognised that the nucleophilic serine, characteristic of the serine β-lactamases is affected by more than just carboxylation of the active site lysine in these OXA enzymes (Drawz and Bonomo 2010; Hamed *et al*., 2013; Bush and Bradford, 2016; Tooke *et al*., 2019). There are currently five separate clades of OXA enzymes with activity against Carbapenems. While four of these (OXA-23, 24/40, 51 and 58) are limited to *A. baumanni*, the fifth, OXA-48 are mobilised on plasmids in Enterobacteriaceae and increasingly associated with carbapenem failure in clinical settings (Poirel *et al*., 2012).

1.2. *Escherichia coli*.

1.2.1. Clinical important of *E. coli*.

Originally identified by Theodor Escherich in childhood diarrhoea in the late 19th century, *E. coli* is a rod shaped, Gram-negative member of the *Enterobacteriaceae* and the most investigated organism and core laboratory workhorse (Blount, 2015). *E. coli* is identified as a common human and animal commensal both transiently passaging through and residing within the gastrointestinal tract and excreted as a proportion of faeces. Subsequently *E. coli* has adapted to life outside of the host and is supported/can be isolated from in/a wide array of environmental niches including sediments and water compartments which in turn further drives *E. coli* evolution through acquisition and exchange of genes from environmental bacteria (Vila *et al.,* 2016; Gonzalez-Alba *et al*., 2019; Touchon *et al.,* 2020).

E. coli can also serve as an opportunistic pathogen in humans (Long *et al*., 2022), and is the most frequent cause of both intestinal pathogenic *E. coli* (InPEC and IPEC), diarrheagenic *E. coli* (DEC) and extraintestinal pathogenic *E. coli* (ExPEC) infections, often resulting in severe bacteraemia and sepsis (Poolman, 2017; Geurtsen *et al*., 2022). As a pathogen, *E. coli* is the causative agent of multiple pathologies such as Haemolytic Uremic Syndrome (HUS) associated with renal dysfunction, to urinary tract infections (UTI) commonly associated with disease of the lower urinary tract (Bonten *et al*., 2021). In the UK, *E. coli* is the most common cause of UTI's (EUSPAUR, 2021).

Virulence determinants feature as a main driver of pathogenicity and can turn a benign strain into a pathogen assisting colonisation and survival in hostile environments. For example, adhesins facilitate 'adhesion' to epithelial cells and promote biofilm formation which in turn may provide protection against a clinical response. Iron acquisition systems facilitate iron scavenging assisting survival and persistence in iron depleted regions such as the urinary tract. While existing as a diverse and potentially critical pathogen the clinical concern surrounding *E. coli* is further compounded with the emergence of drug resistant strains (Vila *et al*., 2016; Johnson *et al*., 2018; Geurtsen *et al*., 2022).

1.2.2. Antimicrobial resistance (AMR) in *E. coli*.

Concern has been growing over the role of pathogenic species belonging to the *Enterobacteriaceae*, notably *Escherichia coli*, in the spread of multi-drug resistance (Carattoli 2009; Nordmann *et al*., 2011; Woodford *et al*., 2011). Clinical treatment of bacteraemia that involve prescribing antibiotics introduce *E. coli* to a variety of stresses that drive the selection for a variety of resistance mechanisms (Woodford and Ellington, 2006; Paitan 2018; Jutkina *et al*., 2018), such as the production of ESBL's as a response to β-lactam exposure (Mederios, 1997).

According to the European Centre for Disease Prevention and Control (ECDC) around 11% of tested *E. coli* isolates in the UK have been reported to be resistant to third generation cephalosporins (ECDC, 2019). A recent 2019 report by Public Health England (PHE) showed an overall 12% increase in Bloodstream infections (BSI's) caused by *E. coli*. According to the yearly report by SONAAR (2017) the rise was 1.4% in Scotland between 2013 and 2017, with the same report identifying a rise in the number of Gram-negative BSI with resistance to 3rd generation cephalosporins. A potentially greater concern is the observed rise in carbapenamase producing bacteria (38% in Scotland 2013 - 2017) which peaked in 2019 but has significantly declined in the following years. However, the figure is observed as only a 2.8% reduction from the preceding 5 years. The most frequently identified genes being OXA-48 like carbapenemases (SONAAR, 2020).

While causing extensive global problems in the clinical environment (Pitout and Laupland, 2008), ESBL carriage has now been shown to shuttle between environmental, clinical and community settings (Canton *et al*., 2008, Burke *et al*., 2012; Gonzalez-Alba *et al*., 2019). This is currently becoming increasingly more relevant to local settings and policy (Drieux *et al*., 2016; Donker *et al*., 2017; Dorado-Garcia *et al*., 2017; Low *et al*., 2019; Parnanen *et al*., 2019; Larsson and Flach, 2022).

1.2.3. Extended-spectrum beta-lactamases (ESBLs) in *E. coli*.

In recent years, the dominant ESBL activity has been with CTX-M enzymes (Pitout *et al*., 2008; Canton *et al*., 2012, Bevan *et al*., 2017) the proliferation of which has been driven by substantial application of third generation cephalosporins. Subsequently, with an increase in

CTX-M β-lactamase mediated resistance, there has been an increase in the application of carbapenems. This has subsequently driven an increase in carbapenem resistant Enterobacteriaceae (CRE) and their dissemination (Hawkey, 2015). CTX-M β-lactamases were first recognised around 1989 in *Enterobacteriaceae*, occurring almost simultaneously in both S. America and Europe and now comprise 172 variants (Bevan *et al*., 2017). The term CTX-M is derived from the preferred hydrolytic activity against cefotaxime (CTX) and M refers to the first isolation in Munich. Subsequently, the first cefotaxime β-Lactamase was referred to as CTX-M1 (Canton *et al*., 2012).

CTX-M β-lactamases were possibly captured and mobilised from chromosomal *Kluyvera* species. The *bla*KLUA genes from *Kluyvera ascorbata* share 100% identity with CTX-M-2 (Humeniuk *et al*., 2002). There are 5 main sub-lineages within the CTX-M β-lactamases, identified as CTX-M groups 1, 2, 8, 9 and 25. A few structural hybrids form different variants the majority of which are possibly derived from CTX-M groups 1 and 9 (Canton *et al*., 2012; D'Andrea *et al*., 2013). CTX-M enzymes have almost displaced other ESBLs such as *bla*TEM and *bla*SHV due to lower fitness costs (Stoesser *et al*., 2016) and being rapidly captured and disseminated by highly mobile genetic platforms within dominant clones that often contain other multiple resistance elements (Canton *et al*., 2012; Johnson *et al*., 2016a). A good example of this would be the ExPEC ST group 131 or B2 025:H4-ST131. ST131 has been shown to be transferable to humans through companion animals (Zhang *et al*., 2018a). While other ST groups such as ST73 and ST127 also have a high occurrence in urinary tract infections (UTI's), they are usually susceptible to extended spectrum cephalosporins. Both ST73 and ST131 are globally recognised, possibly spread by human migration. However, CTX-M genes are commonly found associated with ST131. Subsequently, ST131 is considered as high risk, globally recognised MDR clone predominately associated with community onset infections related to outpatients and geriatrics and linked with increased mortality rates (Petty *et al*., 2014, Dale and Woodford, 2015).

While there are 5 main CTX-M genotype groups, two dominant CTX-M enzyme variants are generally associated with ESBL's, namely CTX-M-14 (group 9) and CTX-M-15 (group 1) reported to have originated in India and China (Bevan *et al*., 2017). These appear ubiquitous throughout all environmental and ecological niches, possibly attributable to anthropogenic influences upon these compartments and subsequent connectivity within and throughout

global regions (Morris *et al*., 2008, Lau *et al*., 2009, Amos *et al*., 2014; Bevan *et al*., 2017; Conte *et al*., 2017).

1.2.4 *E. coli* as an indicator species.

The association of *E. coli* with the digestive tracts of warm-blooded animals and more specifically with human faeces makes it a suitable marker for faecal contamination in aquatic environments (Rice 1991; Ishii and Sadowsky, 2008; WHO 2011;). Standard methods for identification and enumeration of faecal coliforms utilise membrane filtration and selective media utilising β-D galactosidase and β-D glucuronidase biochemical activity (Kaspar *et al*., 1986, Rompre *et al*., 2002; Environment Agency 2009; ISO-9308-1:2014).

1.2.5. ExPEC pathogenic *E. coli*.

As described in section 1.2.2., *E. coli* is the most common cause of ExPEC infections. Multilocus sequence typing (MLST) can identify an *E. coli* strain to a predefined population structure and subsequently infer whether phylogenetic relationships exist between isolates during an outbreak as well as delineate strains within a global context (Zhou *et al*., 2019). Using this typing system, it has been possible to identify specific clonal lineages in *E. coli* as predominant pandemic lineages, namely ST69, ST73 ST95 and ST131 (Riley, 2014). By far the most notorious of these ST's is ST131 which has become synonymous with ESBL dispersal. Stepwise evolution of ST131 was driven by previous interactions with prophages, genomic islands, and insertion sequences, although it appears without ESBL genes notably in clades A & B (Woodford *et al*., 2011, Zakour *et al*., 2016). ST131 clonal sub-group, *fim*H30 or clade C, is the most common lineage evolving from the ST131 clade B. Clade C, while not necessarily driving ST131 expansion (Kallonen *et al*., 2017), maintains a potentially primed phenotype for rapid expansion determined by previous co-evolutionary relationships with plasmid replicons and insertion sequences (IS), is identified as driving the global expansion of CTX-M genes as well as integrating fluoroquinolone resistance. This has now been identified to occur with the two *fimH30* clades, C1/H30-R and C2/H30-Rx (Zakour *et al*., 2016).

Initially, evolution of the more notorious C2 appears to have occurred predominately through IncFII plasmids and IS26 mediated events, with the chromosomal insertion of an MDR gene cassette accompanied by further IS insertions, deletions, and re-arrangements. This

facilitated rapid dissemination due to the amelioration of fitness costs enabled through adaptive plasticity thereby facilitating the retention of beneficial traits (Johnson *et al*., 2016a; Stoesser *et al*., 2016), and with the relationship of these MGE's within C2 strains often unstable, it is possible that plasmid retention is evolving towards chromosomal fixation (Johnson *et al*., 2016a). A 3rd ancestral clade, C0, has been identified with susceptibility to fluoroquinolones and while sharing identity with other clade C members also shares some homology with clade B (no CTX-M carriage) and has been classified as an intermediate strain (Stoesser *et al*., 2016; Zakour *et al*., 2016).

The resistance phenotype carried by C2/H30-Rx provides resistance to fluoroquinolones through mutations in parC and gyrA and resistance to extended spectrum cephalosporins through CTX-M 15 carriage, was a property that first gained ST131 its notoriety (Coque *et al*., 2008; Johnson *et al*., 2013, 2016b; Stoesser *et al*., 2016). More recently the acquisition of the plasmid mediated *aac(6')-Ib-cr* resistant determinant provides resistance against fluoroquinolones and aminoglycosides (Phan *et al*., 2022). C2H30-Rx has even been shown to carry carbapenemases such as OXA-48 (De Toro *et al*., 2017; Ellaby *et al*., 2019) and bla_{KPC2} (Cai *et al*., 2014), highlighting the burgeoning danger that resides within this highly adaptable clonal lineage, epitomized by its high fitness, pathogenicity, adept colonisation capabilities and rapid adaption to hostile heterogenous environmental conditions with a seemingly inherent aptitude for resistance gene acquisition and shuffling (Vila *et al*., 2016).

1.2.6. Horizontal gene transfer (HGT) in *E. coli*.

Horizontal gene transfer is a critical factor in the spread of resistance determinants (Carattoli, 2013). Plasmids enable the bacterial host to adapt to a diverse range of environments occurring through the accumulation of gene combinations that affect virulence, fitness, and resistance. A variety of genetic elements are responsible for carriage and transmission of resistance genes such as genomic islands, plasmids, integrons, integrative conjugative elements, insertion sequences and transposons. Given the ubiquitous nature of such elements, resistance profiles may be complex mosaics of various sequences accumulated via diverse genetic exchanges and interactions between independent mobile elements (Hall *et al*., 2017; Carvalho *et al*., 2015). Plasmids function independently of the host genome and demonstrate a high degree of plasticity facilitated by insertion sequences and their

transposons. Due to the conserved nature of the replication regions, this allows for plasmids to be identified by a process known as incompatibility group typing with each group demonstrating unique characteristics such as copy number and host range (Carrattoli, 2009; Mathers *et al*., 2015).

Plasmids, and in particular IncFII plasmids, are responsible for the transmission of resistance genes such as the CTX-M family genes and recently their prevalence and possible chromosomal integration in 'high-risk' clones such as ST131 CH30 (Mathers *et al*., 2015). Indeed, the plasmid type most associated with *bla*CTX-M-15 gene is the IncFII plasmid group (Carottoli, 2013). The IncFII plasmid has a narrow host range, low copy number and is equipped with addiction and post-segregational killing systems that facilitate maintenance and stability, as occurs in *E. coli* ST131 (Johnson *et al*., 2016a; Pitout and DeVinney, 2017). Importantly, they can be maintained in the absence of selection pressure from antibiotics, indicating a low fitness cost and potential benefit from their carriage. Interestingly, these plasmids can contain multiple replicons that may be either active or silent during replication. IncFII plasmids often capture and carry multiple resistant elements. For example, pEK-499 plasmid identified in an *E. coli* ST131 isolate was identified with *bla*TEM-1, *bla*CTX-M-15, *bla*OXA-1, *aac(6=)-Ib-cr*, *mph*(A), *catB4*, *tetA*, *dfrA7*, *aadA5*, and *sul1* genes (Woodford *et al*., 2009). Further, IncFII plasmids in *E. coli* have been identified with carbapenemase genes such as the *bla*OXA-48 gene (Moussa *et al*., 2020).

To further complicate resistance occurring within clinical settings, is the relative ease that mobile elements flow between different bacteria. In addition, this is shown to occur in a variety of environmental settings and may be promoted by anthropogenic activities (Stokes and Gillings, 2011; Larsson and Flach, 2022). A particular mobile element that epitomises this would be the Class 1 integron (Gillings, 2018). The Class 1 integrons, now seemingly ubiquitous, possess definitive regions of shared ancestral homology, indicating its purveyance to be directly attributable to the commencement of human applied antimicrobial therapy (Gillings, 2014). Their ubiquity is demonstrated in anthropogenically impacted settings and can be intimately tied to faecal contamination. The Class 1 Integrons accumulate a variety of resistance mechanisms in extended arrays of genes called cassettes (Gillings, 2014).

Class 1 integrons can be activated as a bacterial stress response. The genes within the cassette can rearrange and be expressed in a preferential response to an environmental stress. The high replicability rate in response to selection pressures encourages their selection as they confer selection advantages to their bacterial hosts and are autonomous (Gillings, 2014). Due to their association with faeces and, environments influenced by anthropogenically derived selective agents (such as WWTP effluent), are often defined by a high abundance of Class 1 integrons containing multiple resistance genes. Consequently, detection of Class 1 integrons has been used as a proxy for anthropogenic influence on surface water ecosystems (Amos *et al*., 2015, 2017; Gillings, 2018; Quintela-Baluja *et al*., 2021).

1.3. Biological wastewater treatment and the spread of AMR.

Wastewater treatment plants (WWTP's) can provide suitable conditions for human pathogens to act as, or promote, the establishment of reservoirs of resistant determinants (Bengtsson-Palme *et al*., 2017; Suzuki *et al*., 2019). In turn, this may lead to increased opportunities for the persistence of opportunistic human pathogens and/or facilitate horizontal gene transfer of antimicrobial resistance genes to environmental species that may be later transmitted back into opportunistic human pathogens (Leonard *et al*., 2015; O'Flaherty *et al*., 2019). While a reduction of enteric species occurs, the resistant profile, or resistome, of the wastewater community has a strong correlation with the resistome present in the gut microbiome, and more often these genes are not removed during the WWT process. Additionally, microcontaminants that may act as selective agents such as pharmaceuticals, antibiotics, detergents, and metal containing products remain ubiquitous, exposing the microbial communities to stresses that can lead to enrichment and mobilisation of antibiotic resistant determinants (Dealtry *et al*., 2014; Pal *et al*., 2015; Lehmann *et al*., 2016; Ye *et al*., 2017; Burgmann *et al*., 2018; Lu *et al*., 2018; Numberger *et al*., 2019). This effect is not just limited to the various treatment stages, as an association between resistance element abundance and proximity to WWTP outflows has been observed using upstream water quality as a baseline (Amos *et al*., 2014; Li *et al*., 2015; Lekunberri *et al*., 2018). Additionally, there are multiple factors that influence the contribution of WWTP's as environmental polluters, including season, climate, plant design, treatment methods and operational parameters (Amos *et al*., 2015; Gao *et al*., 2018).

Selection processes and bacterial dynamics throughout wastewater networks and the receiving environment appear ecologically compartmentalised and complex (White *et al*., 2016; Auguet *et al*., 2017; Hultman *et al*., 2018; Gao *et al*., 2018; Zhou *et al*., 2018; Quintela-Baluja *et al*., 2019; Paulhaus *et al*., 2019). While underlined by great complexity, WWTPs are currently considered a point source or hotspot for ARG and ARB proliferation and dissemination into the environment (Pruden *et al*., 2012; Czekalski *et al*., 2014; Rowe *et al* 2017; Che *et al*., 2019). Changes that occur in aquatic environments in receipt of wastewater, include enrichment of human and animal pathogen families and an increased abundance of ARG's (Subirats *et al*., 2017; Lekunberri *et al*., 2018; Quintela-Baluja *et al*., 2019). WWTPs influence and legacy effect on the resistome, mobilome as well as species abundance and diversity have been shown to persist into receiving aquatic environments (Corno *et al*., 2019).

Monitoring of AMR in wastewater can provide a cost-effective and unobtrusive method of testing communities and complement existing clinical surveillance data (Marano *et al*., 2020; Robins *et al*., 2022). Recent research suggests monitored changes in microbial populations, specific resistance genes and even chemicals may present an early warning system for potential outbreak of AMR related disease (Riquelme *et al*., 2022). However, the diverse array of microorganisms and potential resistance genes presents a challenge for monitoring. Standardisation of techniques is also required and crucially, must be cost effective, so that data collected can be universally compared across regions (Hassoun-Kheir *et al*, 2021; Klumper *et al*., 2022). The WHO and the Advisory Group on Integrated Surveillance on AMR (AGISAR) recommend the "Tricycle protocol" (WHO, 2021) utilising ESBL producing *Escherichia coli* as a single indicator species for national multisectoral surveillance across human, animal and environmental sectors.

1.4. Study rationale and scope.

Alarm has been growing over the role of Gram-negative bacteria, specifically pathogenic species of the *Enterobacteriaceae*, in the global spread of multidrug resistance and the impact of this on human health (Carattoli 2009; Nordmann *et al*., 2011; Woodford *et al*., 2011). Municipal wastewater and its association to environmental dimension of AMR has recently received increasing attention (Paulshus *et al.,* 2019; Larsson and Flach 2022). The WHO 'One Health' Framework concerning antibiotic resistance recommends unified surveillance of AMR

indicator species (Matheu *et al*., 2017; WHO, 2021). *E. coli* that produce Extended Spectrum Beta-Lactamases (ESBL's) and are resistant to third generation cephalosporins offer a suitable marker for monitoring resistance epidemiology and to better understand the associated risk of AMR in environmental settings (Leonard *et al*., 2018; Hutinel *et al*., 2019; Marano *et al*., 2020). The rationale for this is based on the association of *E. coli* with faecal contamination, and the ubiquity of this species as a human and animal intestinal commensal, colonist, and potential pathogen. Further, ESBL's are responsible for increased disease burden and can cross the species barrier (Matheu *et al*., 2017; WHO, 2021). ESBL's are often found to be ubiquitous in WWTP effluent (Day *et al*., 2019) and have subsequently been detected in receiving water bodies (Korzeniewska and Harnisz, 2013; Brechet *et al*., 2014). Recent studies have identified similarities in the resistance patterns between the clinical and wastewater treatment works (Parnanen *et al*., 2019). Many ESBL producing *E. coli* isolates in the clinic have a strong association with the potentially pathogenic *E. coli* sequence type, ST131, other ST groups which are generally found to harbour CTX-M 15 encoding genes (Coque *et al*., 2008; Lau *et al*., 2008; Dolejska *et al*., 2011; Day *et al*., 2019)

1.5. Aim and Objectives.

The aim of this study is to investigate the presence and diversity of ESBL producing *Escherichia coli* released in tertiary treated effluent to the river Almond from the East Calder Wastewater Treatment Plant. In addition, for selected isolates, to explore the genetic environment of cefotaxime and other antibiotic resistance genes including any associations with mobile genetic elements. In turn, this study will contribute to our understanding of the potential contribution of treated effluent to the freshwater resistome via transmissible antibiotic resistance genes.

The specific objectives of the study are to:

1. Isolate and quantify ESBL producing *E. coli* present in freshwater and river sediment at the outlet of a WWTP using selective isolation techniques.
2. Characterise the phylogenetic diversity of cefotaxime resistant *E. coli* using PCR based assays.

3. Compare the genotypic and phenotypic antibiotic resistance profiles of representative *E. coli* isolates from different phylogroups using PCR based assays and antibiotic sensitivity tests.
4. Establish the local genetic environment of cefotaxime and other antibiotic resistance genes and any associations with mobile genetic elements in selected isolates using whole genome sequencing and bioinformatics tools.

a) b)

Figure 1.2. a) The river Almond catchment area and b) the location of the WWTP at East Calder and the discharge point where treated effluent enters the river Almond.

2. Materials and Methods

2.1. Environmental sampling.

East Calder Wastewater Treatment Works (55°53′57″N , 003°28′08″W) which operates an activated sludge process (ASP) with tertiary treatment for municipal waste was investigated in this study. This large WWTP (capacity 96,000 pe) discharges to the river Almond, near the town of East Calder, West Lothian, Scotland. Treated wastewater effluent samples were collected from the outflow discharge point to the river Almond, on a single sampling day in November 2019. Effluent samples were collected from random points in the outflow channel using a grab sample technique. Briefly, four independent 250ml grab samples were taken at 5 min intervals to give a 1L composite effluent sample. Additional samples were collected from river sediment at random points within three meters from the end of the outflow channel. Approximately 50mm of submerged sediment was collected as 10 samples using a

10mm diameter core sampler. The top 30mm of sediment from each sample was removed from the core sampler and collected in a sterile plastic bag to form a composite sample to be comparatively analysed at a ratio of 1g sediment to 1ml effluent. The samples were transported to the laboratory at 4^{o}C and held at this temperature until use. All effluent and river sediment samples were processed within 24 h.

2.2. Selective isolation and enumeration of non-*E.coli* coliforms and *E. coli*.

The abundance of total NECC (Non-*E. coli* coliforms) and total *E. coli* as well as cefotaxime resistant sub-populations were determined using a protocol adapted from the EPA Method 1603 (2014) and EA Microbiology of Drinking Water (2009) - Part 4. Sediment samples were prepared by mixing 1g of sediment with 9ml PBS and vigorously mixed. Samples were prepared as 10-fold serial dilutions of water and sediment samples down to 10^{-4} in triplicate using phosphate buffered saline (PBS) and 10ml volumes filtered through sterile Whatman filters (0.45µm diam.). Filters were transferred aseptically on to the surface of Coliform Chromselect agar (CCA) plates with or without cefotaxime (4mg L^{-1}). Following incubation at 30^{o}C for 18-24 hours, presumptive NECC (red colonies) and presumptive *E. coli* (blue colonies) were enumerated on those plates containing between 20-80 colonies.

2.3. DNA extraction, PCR, and gel electrophoresis.

Genomic DNA was extracted using a boiled lysate protocol adapted from Cavaco & Hendriksen (2015) with the TE buffer and wash step omitted. Briefly, several colonies from an overnight culture were suspended in 200µl sterile distilled water (dsH_2O) and boiled in a heating block (Flowgen Bioscience) for 10 mins at 95^{o}C. The lysate was placed on ice and centrifuged (Hettich Micro 2000, Germany) at 14,000 rpm for 5 minutes. Deviations from this basic extraction protocol are stated in the relevant PCR assay methods below. Each PCR assay was prepared according to protocols adapted from the relevant literature. PCR amplifications were conducted in a Techne Prime thermocycler (Bibby scientific Ltd, UK).

Amplicons were separated and visualised using 2% w/v agarose gels (Bioline, UK) electrophoresed for 45 minutes at 100 volts, unless otherwise stated. A 100bp or 1Kb ladder (Bioline, UK) was added as a size marker to an appropriate well, as required. Electrophoresed gels were imaged using a Bio-Rad CHEMdoc XRS Molecular imager with Image lab software (Bio-Rad Laboratories Inc.).

2.4. PCR assays.

2.4.1. Simplex PCR for *Escherichia coli* species confirmation.

DNA was extracted and PCR performed as described in section 2.3. Two primer pairs were used to identify presumptive *E. coli* isolates (Table 3.1): a gene specific for *E. coli*, *gad*A*,* which codes for glutamate decarboxylase enzyme (Doumith *et al*., 2012) and the *uid*A gene, which encodes for β-glucuronidase (Berthe *et al.,* 2013). Amplification mixtures contained primers at a final concentration of 1µM for *gad*A and 20µM for *uid*A primers. Amplification of *uid*A was run as a 2 step PCR with the annealing and extension steps combined. The amplification conditions for *uid*A included an initial denaturation step for 3 min at 95°C followed by 25 cycles of 60 s at 94°C and 50°C and a final extension at 72°C for 5 min. The amplification conditions for *gad*A included an initial denaturation for 4 min at 94°C followed by 30 cycles of 30 s at 94°C, 65°C and 72°C and a final extension at 72°C for 5 min. *Escherichia coli* ATCC 47055 was used as a positive control strain for both *gad*A and *uid*A detection and *Klebsiella pneumoniae* (NCIMB 8805) was used as a negative control.

Table 2.1. Primers used for detection of *gad*A and *uidA* marker genes.

Target gene	**Primer sequence (5' – 3')**	**Product length (bp)**
*gad*A	(f) GATGAAATGGCGTTGGCGCAAG (r) GGCGGAAGTCCCAGACGATATCC	373
*uid*A	(f) CATTACGGCAAAGTGTGGGTCAAT (r) CCATCAGCACGTTATCGAATCCTT	600

Abbreviations: f, forward; r, reverse primer.

2.4.2. PCR for Phylogroup assignment.

Genomic DNA extraction and PCR amplification were performed as described in section 2.3. Multiplex PCR was performed according to the method of Clermont *et al*. (2013) using gene specific primers at given concentrations (Table 2.2) to assign *E. coli* isolates to established phylogroups. Modifications to the assay included the replacement of amplification mixture

components to 2x concentration HS MyTaq with nuclease free water. The final primer concentration is given in Table 2.2. The PCR reactions were performed with an initial denaturation step at 94°C for 4 min followed by 30 cycles of 5 sec at 94°C and 20 sec at 59°C (for the quadruplex and group C confirmation), or 20 sec at 57°C (for group E confirmation) with a final extension step at 72°C for 5 min. Interpretation of bands and confirmation assays as described by Clermont *et al.* (2013). Phylogroup assignment was determined according to the presence or absence of bands in the quadruplex assay. Depending upon the sequence of positive or negative results, isolates were either immediately assigned to a phylogroup or alternatively a further two assays were performed to confirm isolates as belonging to either group C or A or group E or D. Two *E. coli* strains, ERI39 and ERI40, were used as in-house positive controls for phylogroups B2 and D, respectively.

2.4.3. Rep-PCR Fingerprinting for strain characterisation.

Rep-PCR was used to amplify the non-coding repetitive extragenic palindrome (Rep) sequences from *E. coli* isolates to generate DNA fingerprint patterns (Malathum *et al.* 1998; Rademaker *et al*., 2008). This allows for identification of clones or near identical populations through unique band patterning. DNA was extracted from overnight cultures grown in Tryptic Soy Broth (Sigma). Briefly, 25µl from each culture broth was centrifuged for 1 min at 12,000rpm and the pellet resuspended in a mixture of 25µl nuclease free water and 75µl of 10% w/v Chelex (Bio-Rad) resin. The suspension was vortexed for 10 s, held at 100^{o}C for 10 min in a heating block (Flowgen Bioscience), vortexed again for 10 s and centrifuged (Hettich Micro 2000) for 3 min at 12,000rpm. PCR assays were performed using primers targeting repetitive extragenic palindromic elements (Table 2.3). Stock primers were diluted to give a final concentration of 50 pmol/µl Table 3.3). The PCR reactions were performed with an initial denaturation step at 95°C for 7 min followed by 35 cycles of 30 sec at 95°C and 60 sec at 40°C and 8 min at 65°C with a final extension step at 65°C for 8 min. Interpretation of the DNA Rep typing pattern was performed according to the methods of Malathum *et al.,* (1998) whereby a difference of 2 or more bands was taken to indicate a different Rep-Type. *Klebsiella pneumoniae* (NCIMB 8805) was used as an internal reference strain to allow comparison of patterns between gels. Electrophoresis was conducted as described in section 2.3 except with a run time of 110 minutes at 70V. To compensate for effects of smiling, 6µl of 1Kb ladder (Bioline, UK) was added to the left- and right-hand lanes.

Table 2.2. Primer sequences and product length used in the PCR assays to determine the phylogroup of *E. coli* isolates (adapted from Clermont *et al.,* 2013).

Table 2.3. Primers used to generate rep-PCR banding patterns (Versalovic *et al*., 1993).

Abbreviations: f, forward; r, reverse primer. I, Inosine; N denote either A,G,T or C nucleotides. Both I and N act as degenerate oligonucleotides.

2.4.4. ESBL resistance gene characterisation (SHV, TEM and OXA).

Genomic DNA extraction was performed as described in section 2.3. Multiplex PCR was performed to detect *bla* genes using primers specific to ESBL encoding genes from the SHV, TEM and OXA families (Dallenne *et al*., 2012) detailed in Table 2.4. The PCR reaction mixes contained *bla*TEM, SHV or OXA primer pairs at a final concentration of 0.4pmol/µl. PCR amplification included an initial denaturation at 94^{o}C for 10 min, followed by 30 cycles of 94^{o}C for 40 s, 60^{o}C for 40 s and 72^{o}C for 1 min, followed by a final extension at 72^{o}C for 7 min. Two *E. coli* strains, ERI39 and ERI40, were used as in-house positive controls for phylogroups.

2.4.5. ESBL resistance gene characterisation (CTX-M group).

DNA was extracted as described in section 3.2. Multiplex PCR amplification was performed using primers specific to ESBL encoding genes for the simultaneous detection of *bla*CTX-M Groups 1, 2, 8, 9 and 25 (Woodford *et al.,* 2006). PCR reaction mixes contained primers (Table 3.5) at a final concentration of 10pmol/µl. Group 8 and 25 share a reverse primer (Gp8/25R) hence the final concentration of this primer was 20 pmol/µl. PCR amplification conditions included an initial denaturation at 94^{o}C for 5min followed by 30 cycles of 94^{o}C for 25 s, 60^{o}C for 40 s and 72^{o}C for 50 s and a final extension at 72^{o}C for 6 min. *Escherichia coli* ATCCC 47055 was used as a negative control. Previously identified cefotaxime resistant isolates, ERI39, AT1.2, AT1.3 and ERI41 were used as positive controls for group 1 and group 9. No controls were available for groups 2, 8 and 25 hence amplicon sizes from Woodford *et al.* (2006) were used to interpret gels.

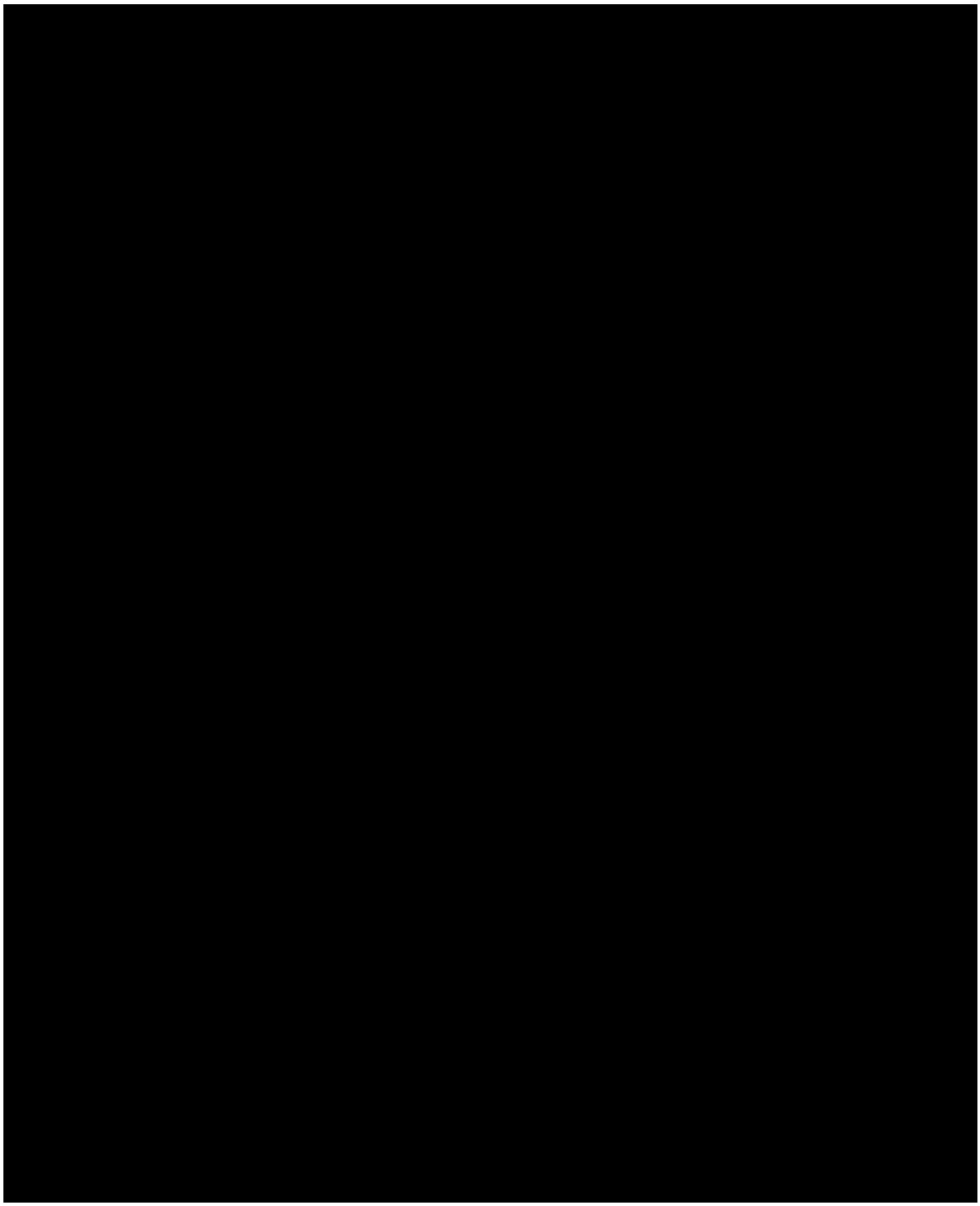
2.4.6.Characterisation of five main *E. coli* sequence types (ST).

DNA was extracted as described in section 3.2. Multiplex PCR was performed using target specific primers for the simultaneous detection of predominant clinically relevant *E. coli* clonal lineages ST69, 73, 95, 131. An additional multiplex PCR was performed using target specific primers for the detection of *E. coli* ST127 (Table 2.6). Adjustments included the replacement of amplification mixture components to 2x concentration HS MyTaq with nuclease free water. Final primer concentrations are given in Table 2.6. Identification of all three gene targets was required for positive confirmation of clonal group ST127. PCR amplification conditions included an initial denaturation at 94^{o}C for 3 min followed by 30 cycles of 94^{o}C for 30 s, 60^{o}C for 30 s and 72^{o}C for 30 s and a final extension at 72^{o}C for 5 min. PCR amplification conditions

Table 2.4. Primer sequences and targets for detection of ESBL (*bla*TEM, *bla*SHV and *bla*OXA) genes (Dallenne *et al.,* 2012).

Table 2.5. Primer sequences used to detect CTX-M groups (1, 2, 8, 9 and 25) and expected product lengths (adapted from Woodford *et al.,* 2006).

Table 2.6. Primers used to detect five predominant sequence types of *E. coli* associated with potential ExPEC pathotypes (adapted from Doumith *et al.,* 2015 and O'Hara *et al*., 2019).



for the ST127 sequence type followed O'Hara *et al*., (2019) with initial denaturation at 98^{o}C for 8 min then 30 cycles of 95^{o}C for 30 s, 58^{o}C for 60 s and 72^{o}C for 40 s and a final extension at 72^{o}C for 5 min.

2.5. Kirby-Bauer disc diffusion method.

The *E. coli* isolates were tested for susceptibility to 12 antibiotics, listed in Table 2.7, using the Kirby-Bauer disc diffusion assay. Freshly prepared Mueller-Hinton (MH: Oxoid CM0337) agar plates (90mm) were acclimatised to room temperature and any plates with excess moisture or surface liquid placed in an incubator 37^{o}C for 1-2 hours until the agar surface was dry. Approximately five colonies were selected from an agar plate incubated overnight using a sterile swab and suspended in 2ml sterile saline solution. This was vortexed to create a homogenous distribution of biomass before adjusting the turbidity to a 0.5 McFarland standard equivalent. The bacterial suspension was then swabbed across the surface of MH plates using a sterile swab before aseptically applying antibiotic discs (Oxoid), six per plate.

The discs were placed at least 24mm apart and within 24mm from the plate edge. Each disc was applied with a small amount of pressure to ensure full and even contact with the agar surface. All tests were conducted in duplicate. The method was conducted using the 15-15-15-minute rule with a maximum of 15 minutes lapse between each step in the procedure (EUCAST, 2021). The plates were inverted and placed in an incubator at 35(±2^{o}C) for 16-18 hours. The diameter of the zones of inhibition was measured to the nearest millimetre on plates with confluent growth using a black backlit background. If the diameter was insufficient the radius was measured and then multiplied by a factor of two to give the diameter. The recorded diameters were interpreted using European Committee on Antimicrobial Susceptibility Testing guidelines for *Enterobacteriaceae* (EUCAST, 2021). Results were recorded as susceptible, intermediate, or resistant. The negative controls used were *E. coli* (ATCCC 47055) and *Klebsiella pneumoniae* (NCIMB 8805). Two previously characterised *E. coli* isolates ERI39 and ERI41 were used as additional positive controls.

Table 2.7. Antibiotics from various classes tested using the Kirby-Bauer disc diffusion method with disc concentration (µg) and diameter of the zone of resistance (mm) according to EUCAST guidelines.

Antibiotic	Disc content (µg)	Zone diameter		Antibiotic class
		S ≥	R <	
Ampicillin	10	14	14	Aminopenicillin
Aztreonam	30	26	21	Monobactam
Cefotaxime	5	20	20	3rd Gen Cephalosporin
Cefoxitin	30	19	19	3rd Gen Cephalosporin
Ceftazidime	10	22	19	3rd Gen Cephalosporin
Chloramphenicol	30	17	17	Chloramphenicol
Ciprofloxacin	5	25	22	Fluoroquinolone
Ertapenem	10	25	25	Carbapenem
Gentamicin	10	17	17	Aminoglycoside
Meropenem	10	22	22	Carbapenem
Tigecycline	15	18	18	3rd Gen. Tetracycline / Glycylcycline
Trimethoprim / Sulfamethoxazole	1.25 / 23.75	14	11	Dihydrofolate Reductase / Sulfonamide

Abbreviations: EUCAST: European Committee on Antimicrobial Susceptibility Testing; Gen.: generation.

2.6. Whole Genome Sequencing of ESBL *E. coli* isolates.

2.6.1 Isolate preparation.

Four *E. coli* isolates, (LIB201, LIB202, LIB210 and LIB211), selected for whole genome sequencing were streaked for single colonies from glycerol (50%, v/v) stocks held at -20^{o}C on to MacConkey agar plates and incubated for 18-24 hours at 24^{o}C. Subsequently, single colonies were used to prepare a lawn of biomass for each isolate on MacConkey agar after incubation for 18 hours at 24^{o}C. The biomass from each plate was collected and resuspended in 10ml PBS buffer and adjusted to OD_{600nm}. Each suspension was centrifuged at 14,000rpm for 3 mins, the supernatant discarded, and the pellet washed twice in 1ml PBS. Each pellet was then resuspended in 0.5ml of 1x DNA shield buffer (Zymo Research, USA) and transferred to a barcoded 2ml screw cap tube, packaged, and shipped according to instruction P650 specifications provided by MicrobesNG UK Birmingham.

2.6.2. *In silico* Genome Analysis utilising Centre for Genomic Epidemiology web-based tools.

Data extension files were received as Prokka output files in various formats . *In silico* genome analysis was first performed using the Centre for Genomic Epidemiology (CGE) based tools (http:/cge.cbs.dtu.dk/services/). Individual isolate sequences were uploaded for analysis in 'Assembled Genomes/Contigs' (FASTA) file format unless otherwise stated. A threshold parameter setting of 90% sequence identity was selected unless otherwise stated, so that a query gene required an identity match of ≥90% to be mapped against a reference gene in the database. In addition, a threshold of 60% minimum coverage length was selected which required any given nucleotide sequence to align or overlap with 60% of a reference gene to be considered a match.

2.6.2.1. Multilocus sequence typing.

MLST finder 2.0.4 (Larsen *et al*., 2012) was used to identify *E. coli* sequence types (ST's) based on Multi-Locus Sequence Typing (MLST). Sequence type is determined using the presence and absence of internal nucleotide sequences in multiple housekeeping genes. The configuration was determined using the *E. coli* #1 Scheme 1 which includes 7 housekeeping genes and identifies 99.6% of relevant alleles (Wirth *et al.,* 2006). Scheme 1 is considered a more complete database than the alternative scheme, *E. coli* #2, which utilises 8 genes and identifies 96.9% of relevant alleles (Jaureguy *et al.,* 2008).

2.6.2.2. *fum*C and *fim*H (CH) typing.

CHTyper (Roer *et al.,* 2018) was used to predict the *E. coli* CH type (Weissman *et al*., 2012). The CH typing scheme allows identification of sub-ST clonal diversity based on *fumC*, one of the household genes used in the seven-locus-based MLST scheme 2, and an internal fragment of the type 1 fimbrial-adhesion-encoding gene *fim*H. A 95% threshold was set as the default for identity to increase the accuracy and enhance the MLST profile prediction (Roer *et al*., 2018).

2.6.2.3. Serotype.

Serotype finder v2.0 (Joensen *et al.,* 2015) was used to identify clinically relevant *E. coli* serotypes based on the unique combination of O-type and H-type antigen genes. The default setting of 85% threshold for identity with 60% minimum length was used.

2.6.2.4. Resistance gene detection.

ResFinder v4.1 (Zankari *et al.,* 2013, 2017; Bortolaia *et al.,* 2020) was used to identify known acquired resistance genes from the curated ResFinder and PointFinder *E. coli* databases. Analysis of known chromosomal point mutations that cause resistance were included in the parameter settings. Additional analysis targeted mutations which conferred no known phenotypic resistance in either ResFinder or PointFinder database. This setting was included as these unspecified mutations may contribute to an antimicrobial resistant phenotype.

2.6.2.5. Mobile Genetic Element.

MobileElementFinder v1.0.3 (Johansson *et al., 2021*) enables the identification of mobile genetic elements (MGE) and plasmids. To enhance detection of the MGE genetic environment within each isolate, antimicrobial resistant and virulence accessory genes were incorporated into the search using ResFinder and VirulenceFinder v2.0 (Camacho *et al*., 2009; Clausen *et al*., 2018) databases.

2.6.2.6. Plasmid replicon types.

PlasmidFinder v2.1 (Carattoli *and Hasman,* 2020) was used to identify and annotate plasmid incompatibility (Inc) groups in each strain. PlasmidFinder provides two database options for analysis, Gram Positive or *Enterobacteriaceae*; the latter was selected using a default of 95% threshold for identity.

2.6.3. *In silico* genome analysis using a bespoke stepwise bioinformatic pipeline.

Further stepwise genome analyses were performed on each isolate using a bioinformatic pipeline developed during this project with web-based genome annotation and visualisation tools. This process involved three key steps as outlined in Figure 2.1: In step 1, contigs with specific Regions of Interest (ROIs) containing genes involved in multiple drug resistance (MDR) and mobile genetic elements (MGE) were identified based on the genome annotations performed using CGE based tools in section 2.6.2. In step 2, these ROIs were resolved further by a stepwise bioinformatic approach using various software to identify additional genes, their length and position on the contig. In step 3, each annotated gene was checked against multiple web-based annotation databases using other comparative tools including Blast and KEGG and a consensus gene map constructed for the ROI's for each isolate.

2.6.3.1. Genome annotation of Regions of Interest (ROI) (Step 1).

The criteria for selection of ROIs in step 1 included identification of CTX-M enzymes and regions containing single or multiple AMR genes resistant to commonly used antibiotics. Additionally, regions that contained MGE's adjacent to an AMR gene identified using CGE data (see section 2.6.2.5) also met the criteria for selection.

WGS were analysed using CARD (Alcock *et al*., 2019) and Prokka (Seeman 2014) annotation tools in Proksee (Grant *et al*., 2008), to identify additional contigs with ROI's. All regions with clustered antimicrobial resistance genes and bla-CTX-M genes with a potential association with mobilisation events due to the presence of plasmids, transposable elements and integrons were identified as potential ROI (Figure 1).

These ROI were further analysed again using PlasmidFinder, MGEfinder and RESfinder, for additional resolution of the genetic environment of the ROI. This included details on gene coordinates and strand direction facilitating initial gene positioning for the ROI during step 3.

2.6.3.2. Comparative analysis (Step 2).

The annotation files provided by MicrobesNG were interrogated using secondary annotation tools provided by RAST v2.0; Rapid Annotation using Subsystem Technology (Aziz *et al*., 2008; Brettin *et al*., 2015) which uses the SEED viewer (Overbeek *et al.,* 2005) as a data visualisation tool. Each isolate genome was additionally reannotated using Prokka hosted in Galaxy. Comparison of genome annotations between Prokka and RAST annotations were

subsequently checked for agreement. Blast (NCBI) was then used to resolve any conflicting annotations between platforms and to provide greater discriminative resolution within a species-specific context. Integrons were identified using SEED viewer and confirmed using IntegronFinder (Cury *et al.,* 2016). Consensus annotation between RAST and Prokka was then contrasted with the data obtained by MGEfinder. Identification of MGE's length, position, and direction, was then inputted to and visualised in Proksee in step 3.

2.6.3.3. ROI visualisation (Step 3).

The construction of a fully annotated consensus map of the genomic environment (gene map) of each ROI was performed using the web-based comparative genomics tool provided by Proksee (Grant *et al.,* 2008). Proksee acts as a surrogate wrapper/host i.e., it imports and runs scripts so that the operator can run multiple programs with no need for proficiency in code. Using the Prokka tool in Proksee, individual genomic features were annotated according to strand direction and position on the contig. The annotation and analysis of amino acid and nucleotide sequences from steps 1 and 2 were then added to the viewer in Proksee. All genomic features were adjusted into the existing track (visible row of consensus annotated genes) or added as a new track. Each visualised genomic feature was colour coded according to representative gene function.

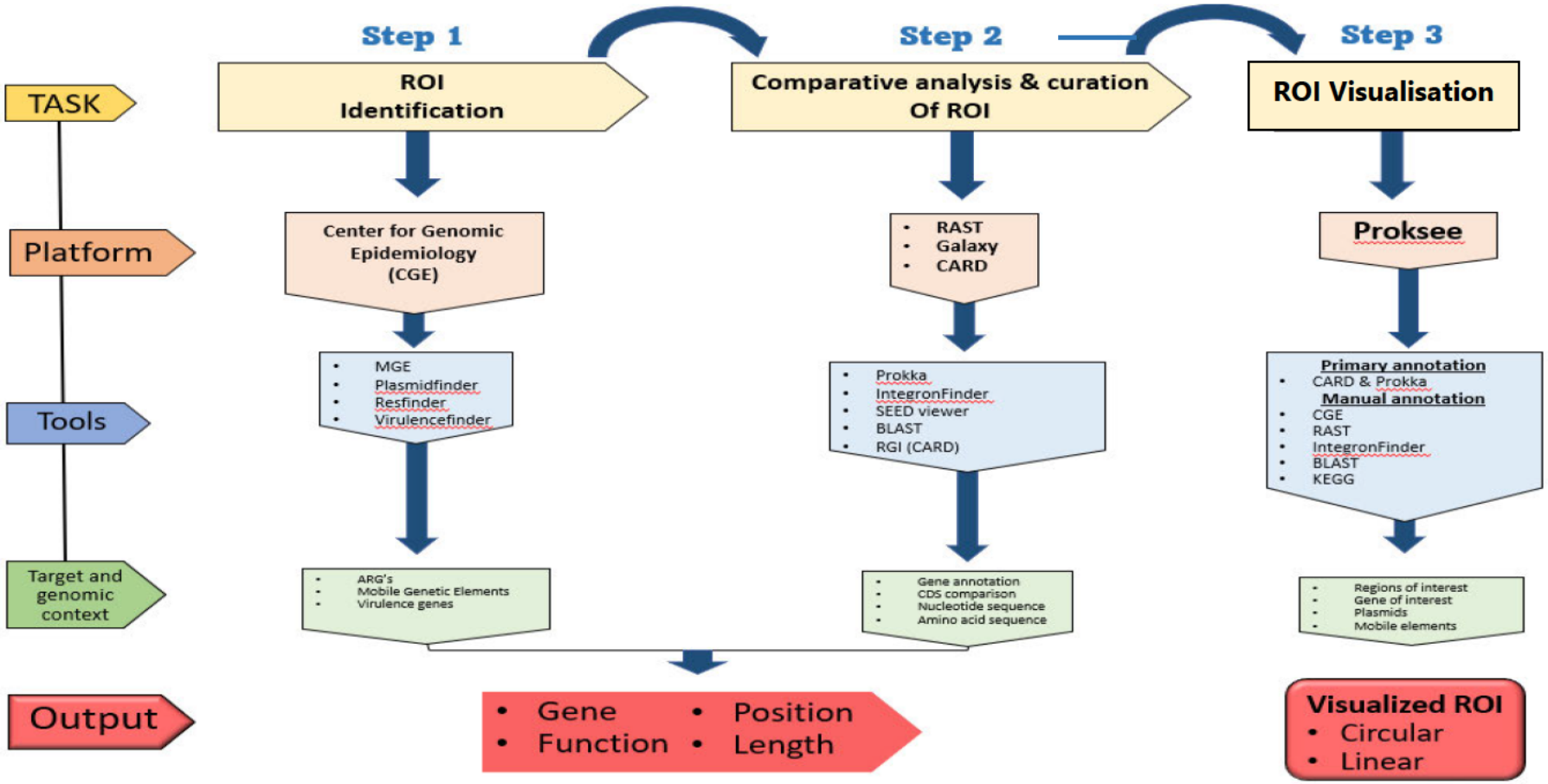


Figure 2.1. Pipeline for comparative analysis of assembled genomes of *Escherichia coli* strains isolated from East Calder WWTP discharge point into the river Almond. Analysis comprised of three steps: Step 1: identification of Regions of Interest (ROI) including MGE's and associated ARG's and Virulence genes, using the Centre of Genomic Epidemiology suite of online tools; Step 2: exploration of the genomic environment within each ROI using various tools to provide annotated data required for comparative analysis and deeper resolution of the genomic environment. Step 3: application of a visualisation tool to facilitate construction of a visual map of the ROI at gene level.

3. Results

3.1. Abundance of *E.coli* and non-*E.coli* coliforms in the WWTP effluent and river sediment.

Total viable counts of presumptive *E. coli* and non-*E.coli* coliforms (NECC) and cefotaxime resistant sub-populations were determined for treated effluent taken at the outflow channel from East Calder WWTP and for adjacent sediment from the river Almond. *E. coli* or NECC colonies recovered on Chromogenic Coliform Agar (CCA) plates were differentiated based on colony colour. Higher numbers of NECC than *E. coli* were recovered from both environmental samples. Total *E. coli* and NECC numbers were higher in the effluent from the WWTP outflow than in the river sediment (Table 3.1). Similarly, the abundance of cefotaxime-resistant *E. coli* and NECC were higher in the outflow effluent (Table 3.1). Cefotaxime resistant *E. coli* accounted for 2.44% of the total *E. coli* count in the effluent whereas none were recovered from the sediment. Cefotaxime resistant NECC were recovered from both outflow effluent and sediment and accounted for 0.84 and 0.13%, respectively (Table 3.1).

3.2. Simplex PCR for *Escherichia coli* species confirmation.

To confirm the specificity of the selective media used in the study (Chromogenic Coliform Agar) and to confirm species identity, 13 isolates (10 *E. coli* and 3 NECC) were selected based on differences in appearance and tested by PCR using two highly specific primer pairs for *E. coli gad*A and *uid*A. The 13 isolates were each assigned a unique reference identification number. No amplicons were detected for the negative control whilst amplicons of approximately 373bp and 600bp were observed for the positive controls confirming the validity of the assay. The results for both the *uid*A and *gad*A PCR assays were in agreement for all 13 isolates screened (Table 3.2). The results revealed that one isolate, namely LIB212 presumptively identified as *E. coli* on CCA medium, was incorrect as no amplicon was detected with either PCR assay (Table 3.2). Additionally, one presumptive NECC isolate (LIB205), identified as a red colony using the CCA, was identified with the *E. coli* specific marker gene.

Table 3.1. Viable counts for total and cefotaxime resistant presumptive *E. coli* and non-*E.coli* coliform (NECC) determined for treated effluent and receiving river sediment collected at the outflow from East Calder WWTP, after membrane filtration and cultivation on Chromogenic Coliform Agar (CCA) with or without cefotaxime (4mg L^{-1}) at 30°C for 18-24 h.

Bacterial group	**Bacterial count (cfu/ml)**	
	Effluent	**Sediment**
Total *E. coli*	8.3×10^3*	8.3×10^1*
CTX-R *E.coli*	2.0×10^2*	0*
Total NECC	2.8×10^5*	4.0×10^4
CTX-R NECC	2.3×10^3*	5.4×10^1*
	Proportion of total count (%)	
Resistant *E.coli*	2.44	0
Resistant NECC	0.84	0.13

*Indicates cfu counts outside of the recommended 20-80 c

Table 3.2. PCR identification of 13 representative isolates presumptively identified by colony colour growing on CCA plates as *E. coli* (blue) or NECC (red) after isolation from WWTP effluent outflow or river sediment.

Isolate code	Sampling point	Presumptive identification	PCR assay	
			***gad*A**	***uid*A**
LIB 201	Effluent	*E. coli*	+	+
LIB 202	Effluent	*E. coli*	+	+
LIB 205	Sediment	NECC	+	+
LIB 206	Sediment	NECC	-	-
LIB 207	Sediment	NECC	-	-
LIB 209	Sediment	*E. coli*	+	+
LIB 210	Effluent	*E. coli*	+	+
LIB 211	Effluent	*E. coli*	+	+
LIB 212	Effluent	*E. coli*	-	-
LIB 213	Effluent	*E. coli*	+	+
LIB 214	Effluent	*E. coli*	+	+
LIB 215	Effluent	*E. coli*	+	+
LIB 216	Effluent	*E. coli*	+	+

Abbreviation: NECC = Non-*E. coli* Coliforms

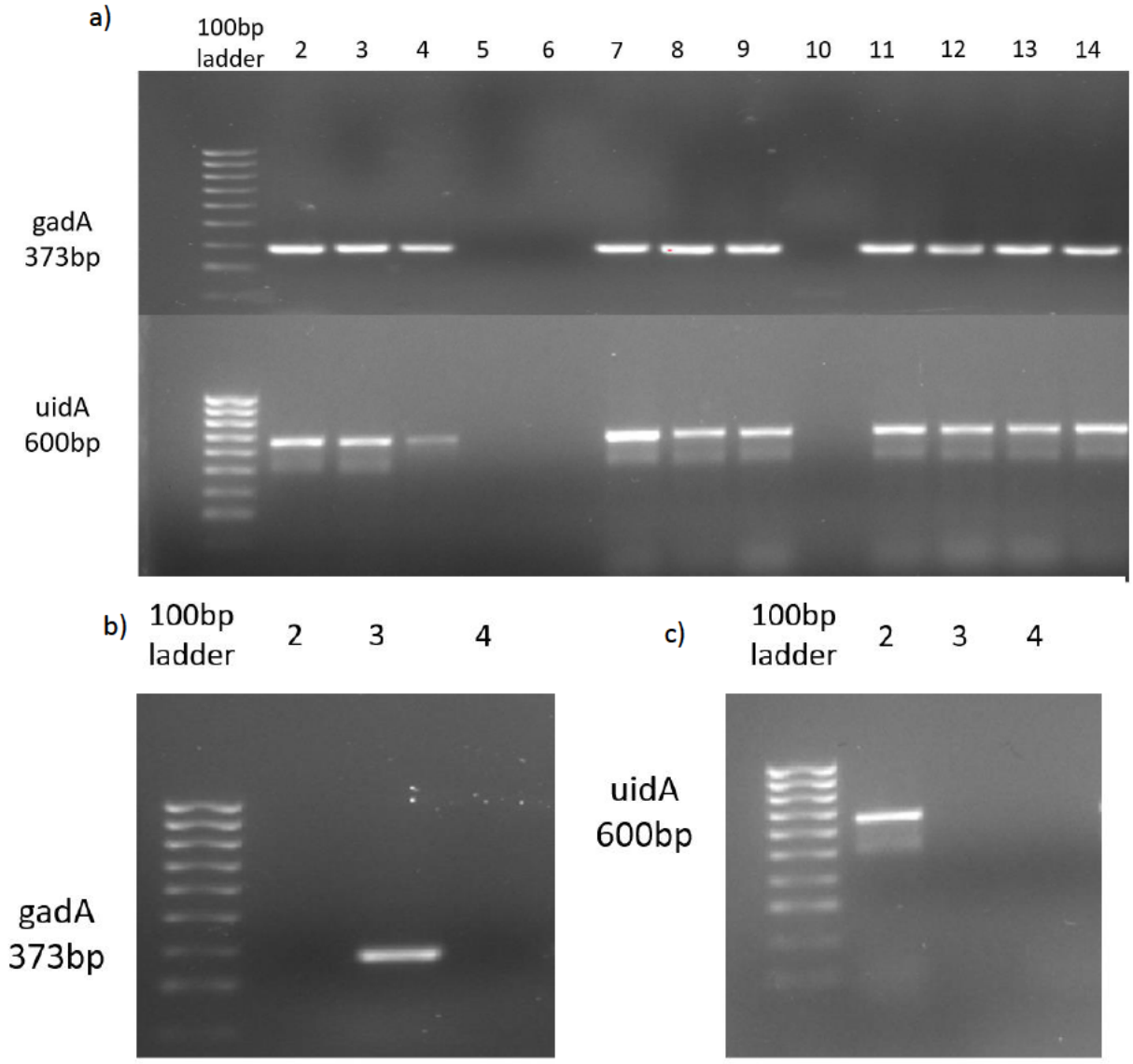


Figure 3.1. Agarose gel (2% w/v) electrophoresis of PCR amplified *gad*A (373bp) and uidA (600bp) gene fragments a) Lanes 1:100bp ladder (bioline), 2: LIB201, 3: LIB202, 4: LIB205, 5: LIB206, 6: LIB207, 7: LIB209, 8: LIB210, 9: LIB211, 10: LIB212, 11: LIB213, 12: LIB214, 13: LIB215, 14: LIB216. b) Lanes 1: 100bp ladder, 2: empty, 3: *E. coli* ATCC47055 (+ control), 4: species NCIMB805 (-control). c) Lanes 1: 100bp ladder, 2: *E. coli* ATCC47055 (+ control), 3: empty, 4: species NCIMB805 (-control).

3.3. Diversity of *E. coli* isolates.

3.3.1 REP-PCR Fingerprinting for strain characterisation.

Due to a technical issue one of the isolates (LIB209) was removed after the species confirmation assay. Subsequently, nine isolates confirmed as *E.coli* by PCR amplification of *gadA* and *uid*A gene fragments were analysed by REP-PCR to investigate their diversity (Table 3.3). Rep-PCR amplifies non-coding repetitive extragenic palindrome (Rep) sequences to generate DNA fingerprint patterns (Malathum *et al.* 1998; Rademaker *et al*., 2008). Differences in genome sequences lead in different band patterns. In this study, isolates with 2 or more band differences were assigned to different REP-types. Five unique REP-PCR banding patterns (labelled REP-types 1 – 5) were generated for the isolates on agarose gels (Fig. 3.2). Two isolates, LIB201 and LIB205, gave rise to identical banding pattern (REP-type 1). Four isolates (LIB 202 and 214-216) were identified as a single REP-type (REP-type 2). The remaining 3 isolates (LIB210, LIB211 and LIB213) were identified as separate REP-types.

3.3.2. PCR for Phylogroup assignment.

A representative isolate from each of the REP-types (n=5) was assigned to phylogenetic groups using the PCR assay of Clermont *et al*. (2013). This assay allows *E. coli* to be categorised into one of eight phylogroups, A, B1, B2, C, D, E, F or G on the basis of surface proteins. Different phylogroups are defined by the presence or absence of amplified gene products (*arp*A, *chu*A, *yja*A and TspE4.C2) identified in different combinations from amplified PCR products. Results in Figure 3.3. (a-c) clearly identified each REP-type as belonging to a unique phylogroup. LIB201 was identified as B2 and indicates that LIB205 also belongs to this phylogroup, LIB 202 was identified as Phylogroup B1 indicating that all REP-type 2 isolates belong to this phylogroup. LIB210 was identified as phylogroup C, LIB211 as phylogroup D and LIB213 as phylogroup F. A confirmation assay distinguished isolates LIB210 and LIB211 from phylogroups A (Figure 3.3b) and E (Figure 3.3c), respectively.

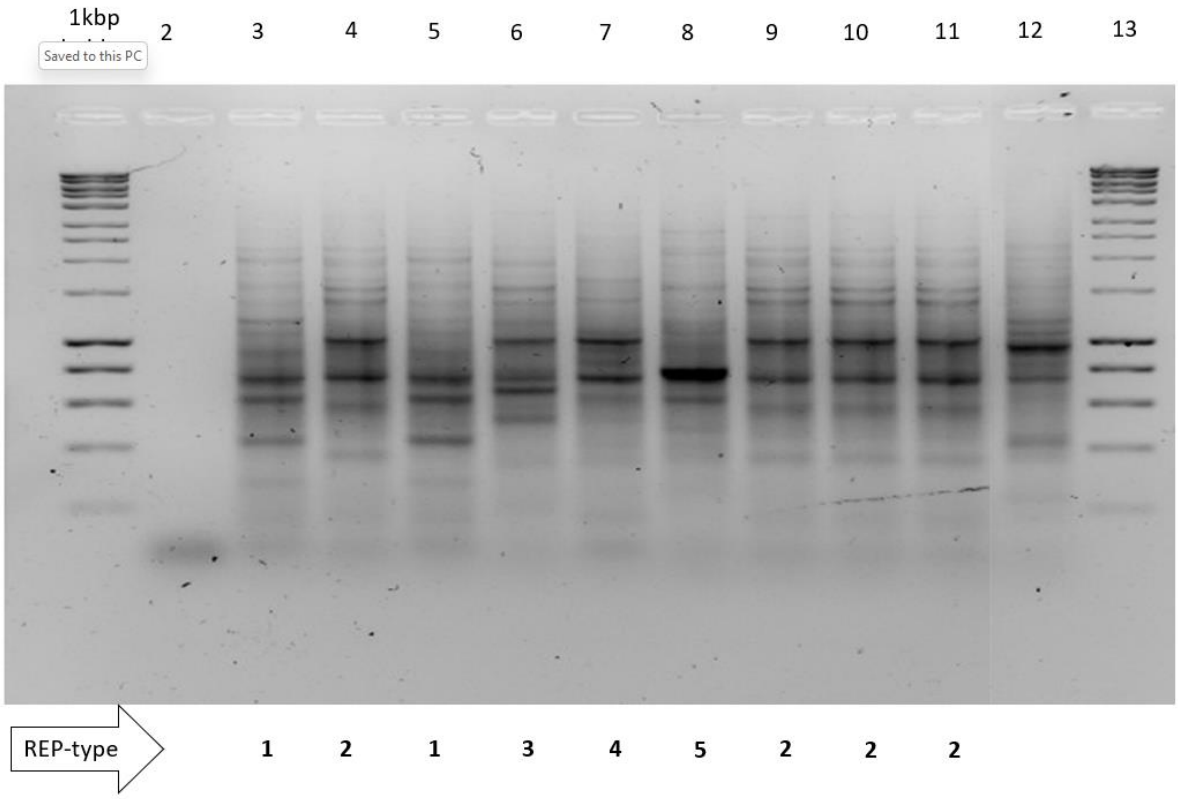
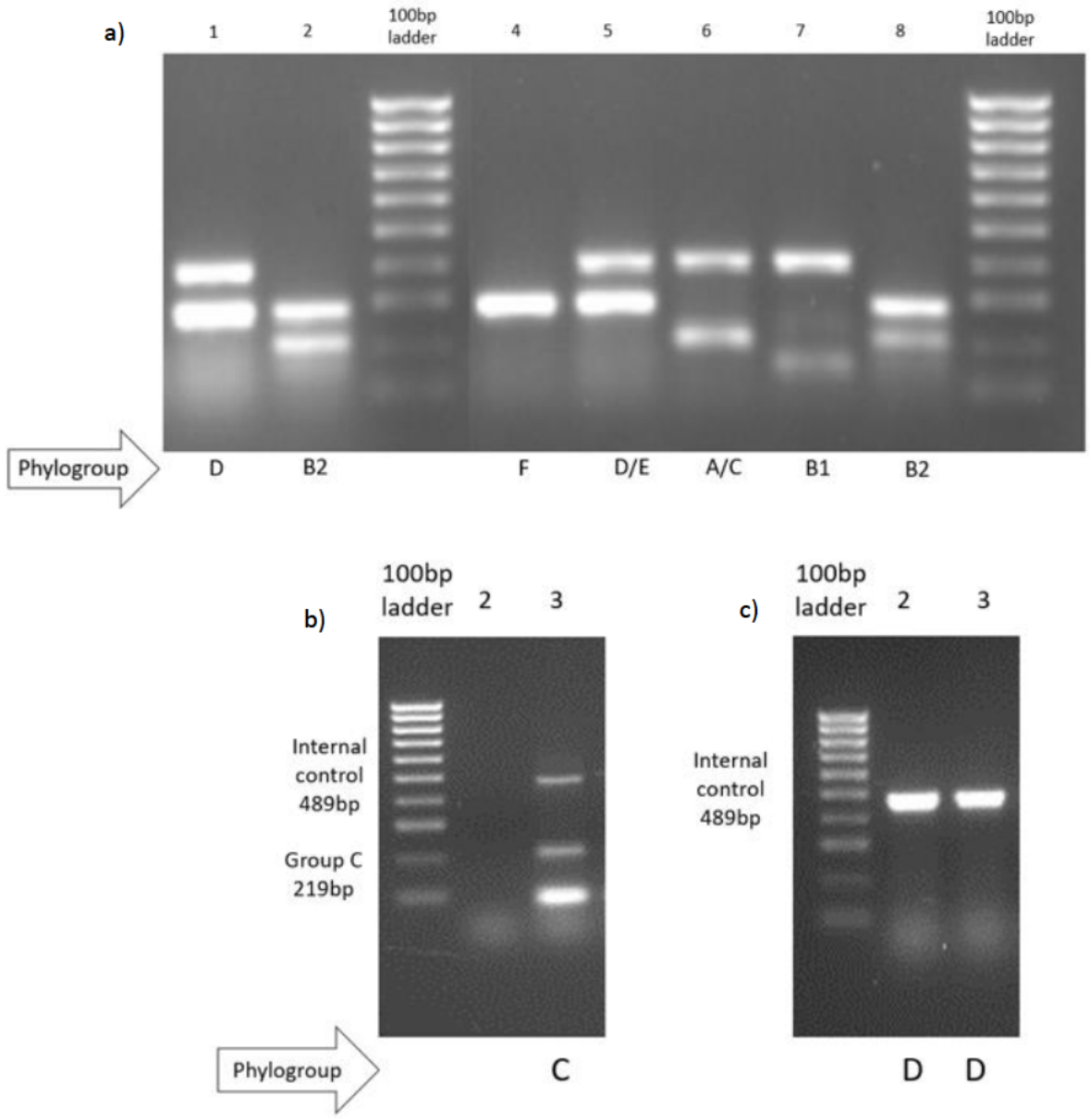


Figure 3.2. Agarose gel image of REP-PCR patterns generated for gadA and uidA confirmed *E.coli* isolates (n=9). Lanes 1 & 15 contain 1kb ladder (source), 2: empty, 3: LIB201, 4: LIB202, lane 5 – LIB205, Lane 6 – LIB210, lane 7 – LIB211, lane 8 LIB213, lane 9 – LIB214, lane 10 – LIB215, lane 11 – LIB216, lane 12 – redundant, lane 13 – redundant, lane 14 – NCIMB8055 used as a reference pattern.

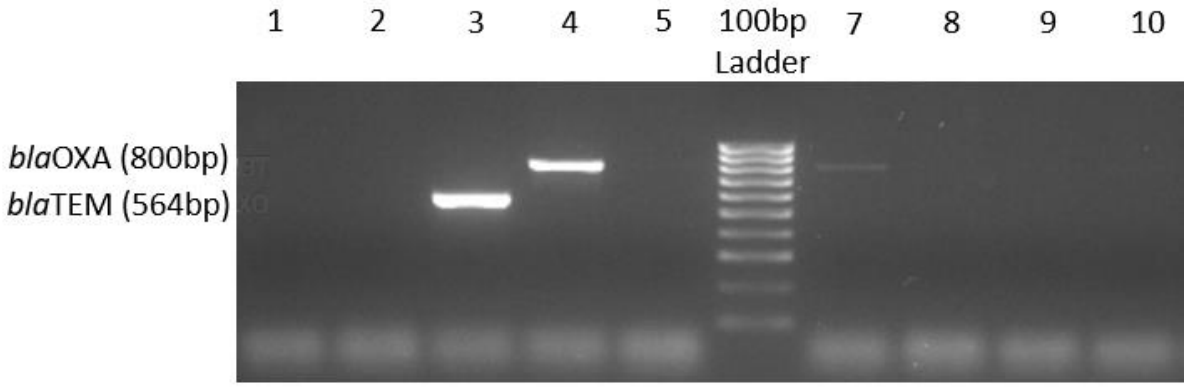


Figures 3.3. Agarose gel images of *E. coli* phylogroup patterns selected as representatives of each unique REP-type (n=5). a) Lane 1 – ERI40 (group D control), Lane 2 – ERI39 (group B2 control), Lane 3 – Ladder (100bp), Lane 4 – LIB213, Lane 5 – LIB211, Lane 6 – LIB210, Lane 7 – LIB202, Lane 8 – LIB201, Lane 9 – Ladder (100bp). b) Confirmation assay required for phylogroup A or C assignment. Lane 1 – Ladder (100bp), Lane 2 – Blank, Lane 3 – LIB210. c) Confirmation assay for phylogroup D or E assignment. Lane 1 – Ladder (100bp), Lane 2 – ERI40 (group D control), Lane 3 – LIB211.

3.4. Characterisation of ESBL resistance.

3.4.1. ESBL resistance gene characterisation (*bla*SHV, TEM and OXA).

Further resistance characterisation of the nine isolates confirmed as *E.coli* by PCR amplification of *gadA* and *uid*A gene fragments was performed to identify *bla*TEM, *bla*SHV and *bla*OXA ESBL marker genes. Only three isolates were identified with these marker genes (Figure 3.4). Two isolates were identified with a *bla*TEM gene (LIB 205 and LIB 211), although the band was weak for LIB211. The remaining isolate, LIB210, was identified with a *bla*OXA gene. *bla*SHV genes were not identified in any of the isolates.



Figures 3.4. Agarose gel image of assay characterising *bla*TEM, *bla*SHV and *bla*OXA ESBL variant presence. Positive identification of these ESBL genes is dependent on the presence of bands with an amplicon size of 800bp for *bla*TEM variants, 713bp for *bla*SHV variants and 564bp for *bla*OXA variants. Lane 1 – LIB201, lane 2 – LIB202, lane 3 – LIB210, lane 4 – LIB211, lane 5 – LIB213, Lane 7 1kb ladder, lane 7 – LIB205, lane 8 – LIB214, lane 9 – LIB215, lane 10 – LIB216.

3.4.2. ESBL resistance gene characterisation (*bla*CTX-M group).

Further resistance characterisation of the nine isolates confirmed as *E.coli* by PCR amplification of *gadA* and *uid*A gene fragments was performed to identify *bla*CTX ESBL marker genes. Only two CTX-M groups were identified; *bla*CTX-M1 genes were identified in all isolates except LIB201 by the presence of a visible band at 415bp. LIB201 was identified with a *bla*CTX-M9 gene by the presence of a visible band at 205bp (Figure 3.5).

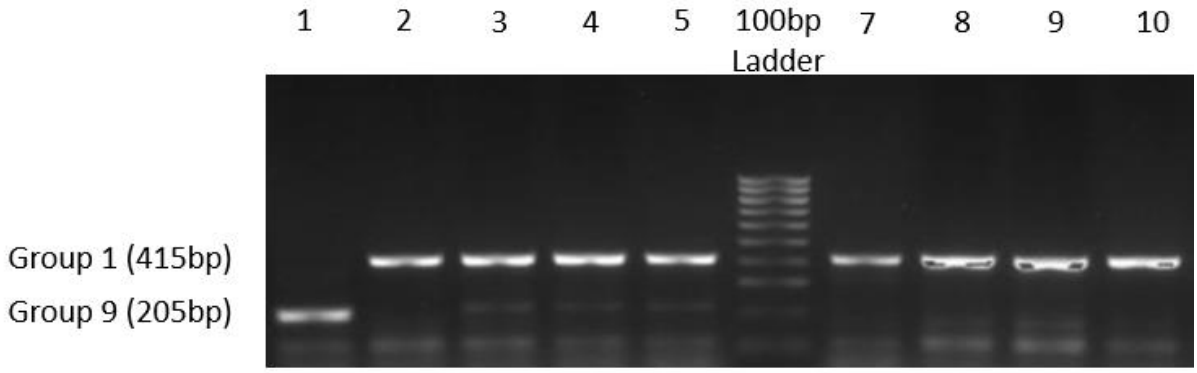


Figure 3.5. Image from gel electrophoresis assay characterising the ESBL activity (CTX-M type) present. Positive identification is dependent on the presence of bands with an amplicon size of 415bp for CTX-M type 1 and an amplicon size of 205bp for CTX-M type 9 an amplicon size. Lane 1 – LIB201, lane 2 – LIB202, lane 3 – LIB210, lane 4 – LIB211, lane 5 – LIB213 // lane 7 – LIB205, lane 8 – LIB214, lane 9 – LIB215, lane 10 – LIB216.

3.5. Characterisation of Five main *E. coli* sequence types (ST).

E. coli isolates LIB201 and LIB211, identified as members of phylogroup B2 and D respectively, were screened by multiplex PCR to determine whether they belonged to extraintestinal pathogenic (ExPEC) sequence type lineages ST69, ST73, ST95, ST131 or ST127. These sequence types only persist in phylogenetic groups B2, and D therefore isolates assigned to other phylogroups were excluded from this assay. Although the in-house ST131 strain (ERI39) performed correctly as a positive control, none of the isolates tested gave rise to an amplicon for any sequence type tested.

3.6. Antimicrobial Susceptibility Testing (AST) for isolates confirmed as *E. coli*.

A selection of 12 discs representing clinically relevant antibiotics (Table 2.7) from 5 classes, commonly used in the treatment of *E. coli* bacteraemia were used to screen for multidrug resistance in the six isolates identified as *E. coli*. These six isolates represent 5 REP-types, including two identified as phylogroup B2 (isolates LIB201 and LIB 205).

Resistance to Cefotaxime was observed in all six isolates and confirmed the phenotype observed during the initial screen using supplementation of 4mg/l Cefotaxime to CCA. Resistance to Ampicillin and Aztreonam was also observed in all six isolates.

Ciprofloxacin and Ceftazidime resistance was observed in five of the six isolates (LIB201, LIB205, LIB210, LIB211 & LIB213). Three of the six isolates (LIB201, LIB205 and LIB210) demonstrated resistance to the trimethoprim/Sulfamethoxazole (SXT) combination disc. Only LIB210 demonstrated resistance to Gentamicin.

All six isolates demonstrated susceptibility to five of the 12 antibiotics (Meropenem, Cefoxitin, Tigecycline, Chloramphenicol and Ertapenem).

3.7 Summary

This study demonstrated that the East Calder WWTP was discharging potentially pathogenic *E. coli* into the river Almond at the time of sampling. A difference in coliform numbers and populations was identified between the effluent and river sediments. This study identified CTX-M genes associated with resistance to 3rd Generation Cephalosporins suggesting ESBL production is occurring in *E. coli* phylogroups regarded as commensal environmental generalists as well as those phylogroups more generally identified as potentially pathogenic strains of *E. coli*. Resistance to multiple antibiotics can be observed to occur in clinically recognised bacteraemia and bacteriuric strains as well as the more generalist strains often associated with animal commensals and the environment.

A considerable proportion of the *E. coli* isolates belonged to the same REP-type and subsequently identified as phylogroup B1. These isolates were isolated from the effluent but also in macrophyte mucilage although these were not included further in the study. Due to the small sample size statistical analysis of sediment and water column diversity was not feasible. However, sediment had significantly lower proportion of cefotaxime resistant and sensitive NECC and *E. coli* in comparison the water column.

Table 3.3. Characterisation of Cefotaxime resistant *E. coli* isolates from treated effluent in the outflow channel from East Calder WWTP and adjacent river sediment within the immediate vicinity of discharge point.

Isolate reference No.	Matrix[a]	*gad*A/uidA[b]	REP type[c]	Phylogroup[d]	ESBL Resistance Gene[e]	CTX-M group[f]	Antibiogram[g]
LIB 201	Outflow	+/+	1	B2	*ND*	9	AMP;CIP;CTX;CAZ; ATM;SXT
LIB 202	Outflow	+/+	2	B1	*ND*	1	AMP;CTX;ATM
LIB 205	Sediment	+/+	1	B2	TEM	1	AMP;CIP;CTX;ATM; SXT
LIB 210	Outflow	+/+	3	C	OXA	1	AMP;CIP;CTX;CAZ; CN;ATM;SXT
LIB 211	Outflow	+/+	4	D	TEM	1	AMP;CIP;CTX;CAZ; ATM
LIB 213	Outflow	+/+	5	F	*ND*	1	AMP;CIP;CTX;CAZ; ATM

[a] Type of matrices at the sampling location point adjacent to East Calder WWTP outflow; [b] Both *gad*A and *uid*A were used as *Escherichia coli* specific gene targets; [c] REP typing was in accordance with Malathum *et al*. (1998) with a difference of 2 or more bands taken to indicate different REP types; [d] Phylogroup interpreted in accordance with Doumith *et al*. (2012) and Clermont *et al*. (2013); [e] β-Lactamase screen for TEM, OXA and SHV groups were used according to Dallenne *et al*. (2010); [f] CTX group assignment was based on the 5 main CTX-M enzyme groups according to Woodford *et al*. (2006); [g] AST was performed comprising of a diverse selection of antibiotic classes. AMP-Ampicillin; CIP – Ciprofloxacin; CTX – Cefotaxime; CAZ -Ceftazidime; ATM – Aztreonam; CN – Gentamicin; SXT - trimethoprim/Sulfamethoxazole

3.8. Whole genome sequence analysis of ESBL resistant *E. coli* isolates.

Whole genome analysis provides an opportunity to further interrogate selected genomes to compliment the previous phenotypic assays performed and provides a deeper discrimination of strain type and resistance profile.

3.8.1. Genome assembly metrics.

Four cefotaxime resistant isolates recovered from the treated effluent from East Calder WWTP and confirmed as *E. coli* underwent whole genome analysis using an external sequencing service (MicrobesNG). These isolates were selected to represent different REP-PCR, phylogroup and serotypes and exhibited different antibiotic resistance profiles. Genome size (bp) ranged from 4.87MB to 5.58MB and the G+C mol% content ranged from 50.57% - 50.82% for the assembled sequences (Table 3.4). The number of coding sequences (CDS) identified on each genome ranged from 4550 on the smallest to 5241 CDS on the largest genome. No relationship could be identified between abundance of resistance genes and genome size. The final genome assemblies contained between 77 to 297 contigs with the shortest contig length at 128bp; the higher abundance of contigs suggestive of regions with lower quality in some assemblies.

Table 3.4. Genome assembly metrics for the ESBL resistant *E. coli* isolates from the effluent outflow at East Calder WWTP.

Isolate code	Total genome size (bp)	G+C content (%)	Total genes (CDS)	Contigs (no.)
LIB201	5027995	50.76	4701	77
LIB202	4876405	50.82	4550	136
LIB210	4940008	50.77	4696	151
LIB211	5583489	50.57	5241	297

Abbreviations: CDS, coding sequence; G+C, guanine + cytosine.

3.8.2. *In silico* genome analysis of *E. coli* isolates

The core genomes of the four isolates were analysed using three *in silico* strain tying methods according to MLST, which facilitates identification of potential pathogens through strain typing; CHtype, which expands on MLST data to evaluate sub-ST clonal diversity; and serotype which can be used to identify potentially pathogenic strains and classify them into clonal

groups (Table 3.5). The four isolates were assigned to different MLST types (Table 3.5); two of these sequence types (ST131 & 69) are commonly identified in human and animal bacteraemia, and another, ST44 which resides within the ST10 clonal complex, is linked to pathogenesis. A total of 4 *fim*H types were identified with *fim*H54 being identified in 2 isolates (LIB202 and LIB210) whilst four separate *fum*C genes were observed (Table 3.5). The isolate all had different O-antigen types.

Table 3.5. Multi-locus sequence, Sero- and CH-types determined by *in silico* analysis of the assembled genomes of ESBL resistant *E. coli* isolates from the effluent outflow at East Calder WWTP.

Isolate code	MLST	CH Type	Serotype
LIB201	ST131	H30-C40	O25b:H4
LIB202	ST58	H54-C4	O120:H25
LIB210	ST44 (CC10)	H54-C11	O101:H4
LIB211	ST69	H27-C35	O15:H18

Abbreviations: CC, clonal complex; CH, FumC and FimH type; MLST, multi-locus sequence type; ST, Sequence type

3.8.3. Antimicrobial resistance genes identified on the *E. coli* genomes by *in silico* analysis.

The genome sequence data provided by WGS was interrogated using a bioinformatics pipeline developed in-house with multiple web-based open-source bioinformatic tools. Resistance and resistance gene mobility in the genome of each *E. coli* isolate was characterised and visualised as outlined in Fig 2. in the Materials & Methods section. The genomes of all four isolates were found to harbour between 1 (LIB202) and 14 (LIB210) resistance genes (Figure 3.6) belonging to four or more antibiotic classes (Table 3.6) Isolates with a wider abundance of phenotypic resistance (LIB201, 210 and 211), featured a greater diversity in predicted resistance genes using RESFinder.

Genes encoding ESBLs (*bla*CTX genes), and other β-lactams were found to be present on the genome of all four isolates LIB201, LIB202, LIB210, LIB211. Identification of *bla*CTX, *bla*TEM and *bla*OXA genes confirmed the findings from PCR assays and confirmed the phenotypic resistance to β-lactams and susceptibility to cefoxitin observed in the antibiogram. Importantly, these genes were found on contigs identified as plasmids and/or associated with mobile elements such as insertion sequences, transposable elements and/or integrons and in many instances alongside other antimicrobial resistance genes.

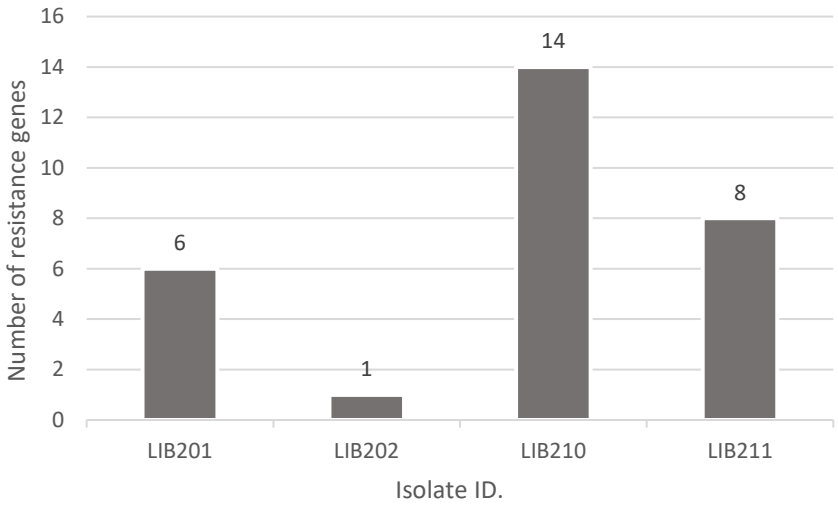


Figure 3.6. Abundance of antimicrobial resistance genes present on assembled genomes of the ESBL resistant *E. coli* isolates

Both isolates LIB201 and LIB202 were identified with only one ESBL gene. The *bla*CTX-M27 in LIB201 with predicted resistance to four β-lactams and the *bla*CTX-M1 gene in LIB202 with predicted resistance to nine beta lactams. The presence of these two genes explained the phenotypic resistance to Ampicillin, Cefotaxime and Aztreonam (AMP;CTX&ATM) identified in the antibiogram (Table 3.6) in these isolates. However, resistance to ceftazidime was predicted in LIB202 but no phenotypic resistance was observed in the antibiogram. All other *bla*CTX-M genes were predicted to confer resistance to ceftazidime and confirmed the results of the antibiogram in all isolates other than LIB202.

Both LIB210 and LIB211 were identified with a *bla*CTX-M15 gene. This *bla*CTX-M15 gene gave predicted resistance to nine β-lactams. This explains the resistance to Ampicillin, Cefotaxime and Aztreonam (AMP;CTX&ATM) identified in the antibiogram (Table 3.6). In both these isolates, additional β-lactamase genes were observed with a *bla*OXA-1 in LIB210 and *bla*TEM-1B in LIB211.

The *bla*OXA-1 gene in LIB210 was predicted to confer resistance to four of the same β-lactams as the *bla*CTX-M15. The *bla*OXA-1 gene gave predicted resistance to β-lactam antibiotic/inhibitor combinations and was the only isolate with this type of predicted

resistance. The *bla*TEM-1B in LIB211 was predicted to confer resistance to four of the same β-lactams as the *bla*CTX-M15 gene as well as cephalothin.

However, although marginal (1 nucleotide over and 99.88% identity), the *bla*OXA-1 gene was identified as a non-perfect match by RESfinder and subsequently it could not be confirmed whether this non perfect score affected the efficacy of the gene as no phenotypic screen was performed for this enzyme beyond those that could be explained by the presence of the *bla*CTX-M15 enzyme.

Quinolone resistance was detected in three isolates with the identification of resistance genes, *aac*(6')-Ib-cr gene in LIB210 and *qnr*S1 in LIB211. Mutations in the *gyr*A genes at positions p.S83L and p.D87N were also identified in two isolates LIB201 and LIB210. The mutations in the *gyr*A genes were the only mechanism of resistance observed in LIB201. As shown in Table 3.6, LIB210 carried an additional mechanism for quinolone resistance, an *aac(6')-Ib-cr* gene. This gene indicated predicted resistance to both Quinolones and Aminoglycosides and was only identified in isolate LIB210.

The *aac(6')-Ib-cr* gene was the only quinolone resistant mechanism with predicted resistance to fluoroquinolone identifying LIB210 to carry resistance to three different quinolone antimicrobials. In contrast mutations in the *gyr*A genes in LIB201 conferred resistance to both ciprofloxacin and nalidixic acid and the *qnr*S1 gene in LIB211, only conferred resistance to a single quinolone, ciprofloxacin. In summary, quinolone resistance identified using ResFinder confirmed the phenotypic resistance to ciprofloxacin in the antibiogram.

Genes identified as *qac*E which facilitates resistance to quaternary ammonium compounds and *sit*ABCD which provides resistance to hydrogen peroxide, were found in three isolates but did not meet the minimum criteria for sequence identity and sequence length, respectively (Table 3.6). Being that phenotypic resistance to these antimicrobials was not screened for no confirmation could be assumed.

Three folate pathway agonists were identified amongst three isolates (LIB201, LIB210 and LIB211). ResFinder identified multiple genes that could be responsible for the phenotypic resistance. Sulphonamide resistance genes were identified in the same three isolates, represented by *sul*1 gene in each isolate, and an additional *sul*2 gene was identified in LIB201. Moreover, predicted resistance to Trimethoprim was only identified in two of these three

isolates (LIB201 and LIB210) with identification of a *dfr*A17 gene. Being that the only folate pathway agonist identified in LIB211 was *sul*1 gene it can be suggested that this was the cause of phenotypic resistance observed in all three isolates. Folate pathway agonists were screened for using a combination disc of trimethoprim and sulphonamide and ResFinder confirmed the results from the antibiogram.

Genes conferring resistance to phenolics were not identified amongst the isolates apart for isolate LIB210 which was found to harbour a gene most closely matched to *cat*B3, a chloramphenicol acetyltransferase gene using RESFinder, although this identification did not meet the 90% similarity threshold. The results of the antibiogram did not reveal resistance to chloramphenicol in LIB210 or any of the *E. coli* isolates.

Five genes involved in resistance to aminoglycosides *aac(6')-Ib-cr, aac*(3)-lla, *aph*(3)-lb, *aph*(6)-ld and *aad*A5 were identified in total (Table 3.6). Isolate LIB201 had three of these, *aph*(3)-lb, *aph*(6)-ld and *aad*A5 which all predicted resistance to streptomycin. The *aad*A5 gene was also predicted with resistance to spectinomycin. This gene was the only aminoglycoside resistance gene identified in isolate LIB211 and similarly had predicted resistance to both streptomycin and spectinomycin.

By far the most diverse predicted resistance to aminoglycosides was observed in LIB210. Similar to isolate LIB201, three different aminoglycoside resistance genes were identified in LIB210. However, while both these isolates carried *aad*A5 genes LIB210 carried *aac(6')-Ib-cr* and *aac*(3)-lla genes instead of the *aph*(3)-lb and *aph*(6)-ld genes carried by LIB201. This provided predicted resistance to an additional six aminoglycosides. The only aminoglycoside screened for in the antibiogram was gentamicin and resistance to this was identified in isolate LIB210. ResFinder confirmed this resistance to be due to the *aac*(3)-*IIa* gene.

Table 3.6. Types of resistance genes identified using ResFinder on the assembled genomes of *Escherichia coli* isolates from the effluent outflow at East Calder WWTP. Antibiogram results are included in the bottom row.

Resistance to antibiotics	***E. coli* isolate**			
	LIB201	**LIB202**	**LIB210**	**LIB211**
Antibiotic class:				
Aminoglycoside	*aad*A5; *aph*(3")-Ib *aph*(6)-*Id*	ND	*aadA5* *aac*(3)-*IIa*; *aac*(6')-*Ib-cr*;	*aad*A5
Beta-lactam	*bla*CTX-M27	*bla*CTX-M1	*bla*CTX-M15; *bla*OXA-1[b]	*bla*CTX-M15; *bla*TEM-1B
Disinfectant	*sit*ABCD[a]; *qac*E[b]	ND	*sit*ABCD[a]; *qac*E[b]	*sit*ABCD[a]; *qac*E[b]
Folate Pathway Agonist	*dfr*A17; *sul*1; *sul*2	ND	*dfr*A17. sul1	*sul*1
Quinolone	(see mutations)	ND	*aac*(6')-*Ib-cr*; (see mutations)	*qnr*S1
Macrolide	*mph*(A)	ND	*mph*(A)	*mph*(A)
Phenicol	ND	ND	*cat*B3[b]	ND
Tetracycline	*tet*(A)	ND	*tet*(B)	ND
Mutations	*gyr*A; *par*C; *par*E	ND	*gyr*A; *par*C; *par*E	ND
Antibiogram	AMP;CIP;CTX;CAZ;ATM; SXT	AMP;CTX; ATM	AMP;CIP;CTX;CAZ;CN; ATM; SXT	AMP;CIP;CTX;CAZ;ATM

Abbreviations: AMP, ampicillin; ATM, Aztreonam; CAZ, Ceftazidime; CIP, Ciprofloxacin; CN, Gentamicin; CTX, Cefotaxime; SXT, Trimethoprim-Sulfamethoxazole.

[a], non-perfect match (<threshold of 60% minimum length). [b], non-perfect match (<threshold 90% identity). ND, none detected.

Table 3.7. Mobile genetic elements and associated resistance genes identified using MobileElementFinder on the assembled genomes of *Escherichia coli* isolates recovered from the effluent outflow at East Calder WWTP.

Genotypic characteristic	***E. coli* isolate**			
	LIB201	**LIB202**	**LIB210**	**LIB211**
Plasmid no.	6	1	8	8
Resistant plasmids	IncFII	Incl1	IncFII (n=2) IncFIB/IncFIA	IncFII
Composite transposon and insertion sequence associated resistance	**(IS6)** CTX-M27	**---(ISAs1)** mdf(A)	**---(IS6)** *aac*(6')-Ib-cr, *bla*OXA-1, catB3	**(IS6*cn*)-[ISKra4*is*])** *qnr*S1, *bla*CTX-M15]
Plasmid mobilised resistance	**IncFII** **[IS6*cn*- CTX-M27] –** sul2, *aph*(6)-Id, tet(A) and *aph*(3")-Ib	Incl1 CTX-M1	IncFII (n=2) **[IS6*cn*- *aac*(6')-Ib-cr,** ***bla*OXA-1, catB3]** --- *aac*(3)-IIa, *bla*CTX-M15 IncFIB/IncFIA *sit*ABCD	IncFII *aad*A5, *qac*E, sul1, mph(A) **–[IS6*cn*-(*ISKra4*) – *qnr*S1, *bla*CTX-M15]**
Integrons and gene cassette	**Int1** *dfr*A17 *aad*A5 *qac*E sul1	ND	**Int1** *dfr*A17 *aad*A5 *qac*E sul1	ND

Abbreviations: Cn/*cn* & IS/*is* = composite transposon with an embedded insertion sequence. --- = separate entries. ___ = plasmids

Although predicted resistance to tetracyclines was identified by RESFinder, with *tet*(A) and *tet*(B) genes identified in isolates LIB201 and LIB210, no predicted resistance to last line tetracyclines were identified in ResFinder. Resistance to Tigecycline, was screened for in the antibiogram but no phenotypic resistance to Glycylcycines (tetracycline derivatives) was observed. Last line carbapenem antibiotics ertapenem and meropenem were screened for in the antibiogram but no isolates presented phenotypic resistance. This observation was confirmed by the ResFinder analysis as no genes were identified with predicted resistance to these antibiotics.

Although no macrolides were screened for in the antibiogram a mph(A) gene was identified by ResFinder in three isolates, LIB210, LIB210 and LIB211 and in all cases indicated a predicted resistance to the same four macrolides.

3.8.4. Mobile genetic elements identified on the genomes of the *E. coli* isolates.

To further characterise ESBL resistance the MGEFinder tool was used to identify contigs that contained regions of interest (ROI's) with potential mobile resistant elements (plasmids and composite transposons) and utilise this information as a baseline for annotation. Subsequent analysis of these ROI's was conducted using the annotation tools provided by web-based platforms annotation tools provided by RAST and Galaxy to further visualise and compare the genetic environment of these MGE's.

3.8.4.1. Plasmids.

Multiple plasmids were detected in each of the cefotaxime resistant *E. coli* isolates except for isolate LIB202 which carried only one plasmid (Table 3.7). Although other IncF plasmid types (FIA/B) were identified in LIB210, the majority of resistance genes identified in three isolates (LIB201, LIB210 and LIB211), were most associated with Incompatibility F group (IncFII) type plasmids. The exception, isolate LIB202, instead carried the one resistance gene detected, *bla*CTX-M1, on an incompatibility group I1 (IncI1) plasmid (Table 3.7).

The IncFII plasmids found in isolates LIB201, LIB210 and LIB211, carried ESBL's (*bla*CTX-M15, *bla*CTX-M27 and *bla*OXA), which were commonly associated with composite transposons. However, the *bla*CTX-M15 resistance gene identified in isolate LIB210 was slightly upstream from the transposon which carried the *bla*OXA enzyme. The only resistance gene not associated with a plasmid was the *bla*TEM-1B gene found in isolate LIB211.

On plasmids, resistance genes not carried by transposons appeared as “resistance gene-rich” regions adjacent to the composite transposons. However, two “resistance gene-rich” regions containing folP/sul2, *qac*E_Δ1, ANT(3’’)-Ia and *dfr*A17 genes adjacent to each other were identified inn isolates LIB201 and LIB210. These both featured on separate small contigs with no association to any plasmids. While these two “resistance gene-rich” regions in isolates LIB201 and LIB210 were identical to each other (Figures 3.9 and 3.10), a similar pattern with just the folP/sul2, *qac*E_Δ1, ANT(3’’)-Ia genes adjacent to each other did occur on the IncFII plasmid supported in LIB211 (Figure 3.8).

3.8.4.2. Composite transposons.

Both insertion sequences and their composite transposons are known to have a significant association with ARG’s. This association can be further resolved to specific IS family members especially IS6 family of IS and includes the clinically relevant IS26.

Insertion sequences were frequently identified by MGEfinder in all isolates. Three IS6 composite transposons frequently associated with resistance gene mobilisation were identified along with the associated IS6 insertion sequences and were commonly identified as carrying resistance within plasmids (Table 3.7).

These were identified in isolate LIB201 with a *bla*CTX-M27 gene (Figure 3.6) and in isolate LIB210 incorporating *aac(6’)-Ib-cr*, two *bla*OXA-1 and *cat*B3 resistance genes (Figure 3.7). A third IS6 composite transposon was identified in isolate LIB211. Although flanked by IS6 insertion sequences the composite transposon had numerous transposases within it including an ISKra4 family member insertion sequence. This latter transposon carried *bla*CTX-M15 and *qnr*S1r resistance genes and are visible as red between positions 31563-37735 in Figure 3.8.

3.8.5. Integrons identified on the genomes of the *E. coli* isolates.

Further analysis of these ROI using the SEED viewer in RAST, and IntegronFinder in Galaxy, broadened the search to include Integrons. Integrons are usually associated with chromosomes and feature 3 core elements that form a gene capture system capable of recruiting genes into expansive pools of cassettes. No integrons were detected on the genomes of the *E. coli* isolates using MGEFinder however, further analysis with RAST and IntegronFinder identified two Int1 integrase genes in isolates LIB201 and LIB210 (Table 3.7). The integrons identified on isolates LIB201 and LIB210 were identical and formed the

“resistance gene-rich” region identified by MGEfinder (section 3.8.4.1), now identified as multiple drug resistant gene cassettes containing, folP/sul2, *qac*E_Δ1, ANT(3’’)-Ia and *dfr*A17 (Figure 3.9 and 3.10, respectively).

Upon further investigation of a specific region of these integrons in both LIB201 and LIB210 (Figures 3.9. and 3.10), using the online BLASTn tool, a highly conserved region was identified between the attC binding site and *fol*P gene. These were identified in another study by Qunitela-Baluja *et al*., (2021), who identified that these two genes are fused together and as such can be identified by this highly conserved region. This identifies both these integrons as Class 1 and highly specific to anthropogenic polluted environments and not clinical settings. Although not identified by MGEfinder, the RAST annotation tool identified additional genes associated with the macrolide resistant region which may have an association with an IS6 family insertion sequence upstream from the cassette. This macrolide resistance region upstream from the integron was absent in LIB210.

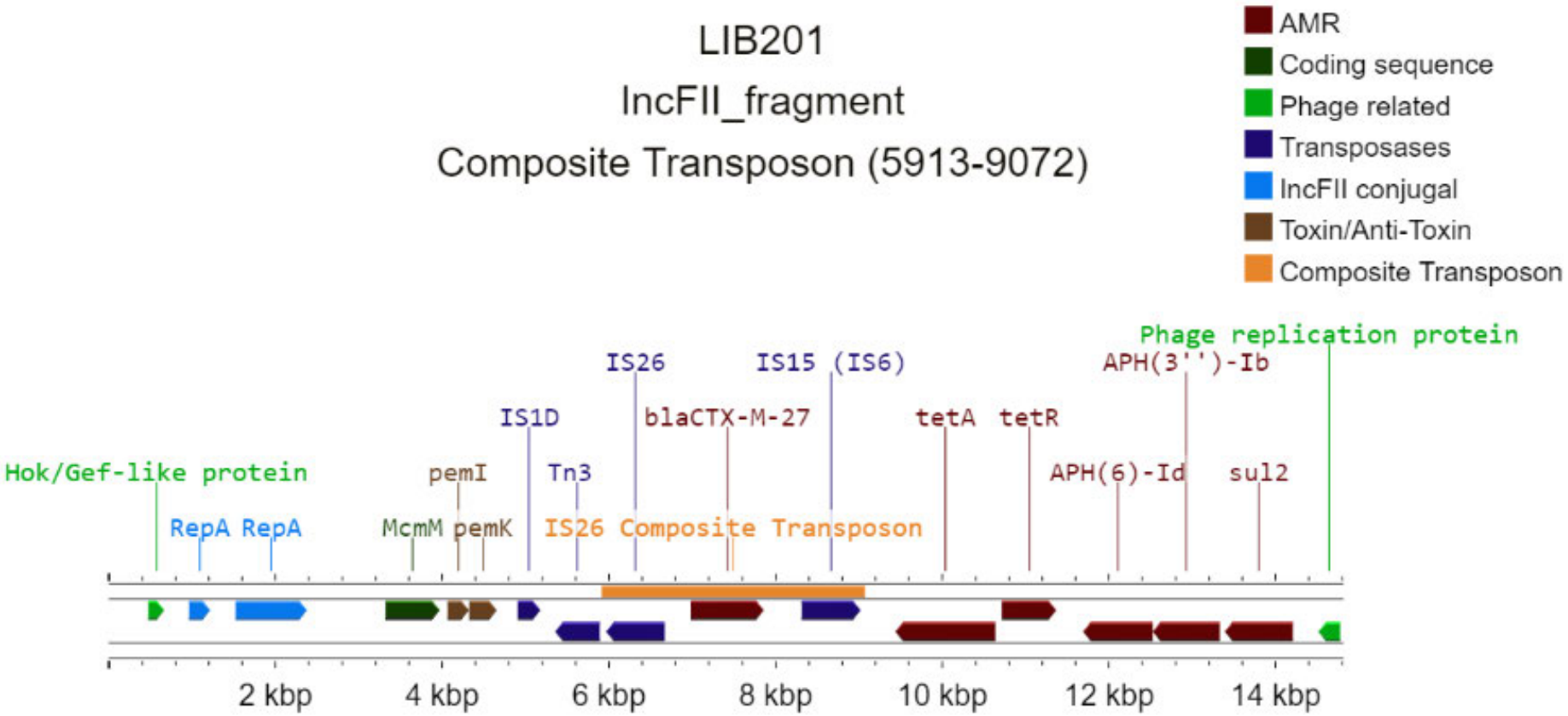


Figure 3.7. Linear map of the genome section in *E. coli* isolate LIB201 that encompasses the genetic environment of the *bla*CTX-M27 gene. Putative genes are coloured according to the predicted functions of their products. The orange region (position 5913-9072) is an IS26 composite transposon identified by MGEFinder in CGE. The direction of the arrows indicates the orientation of gene transcription.

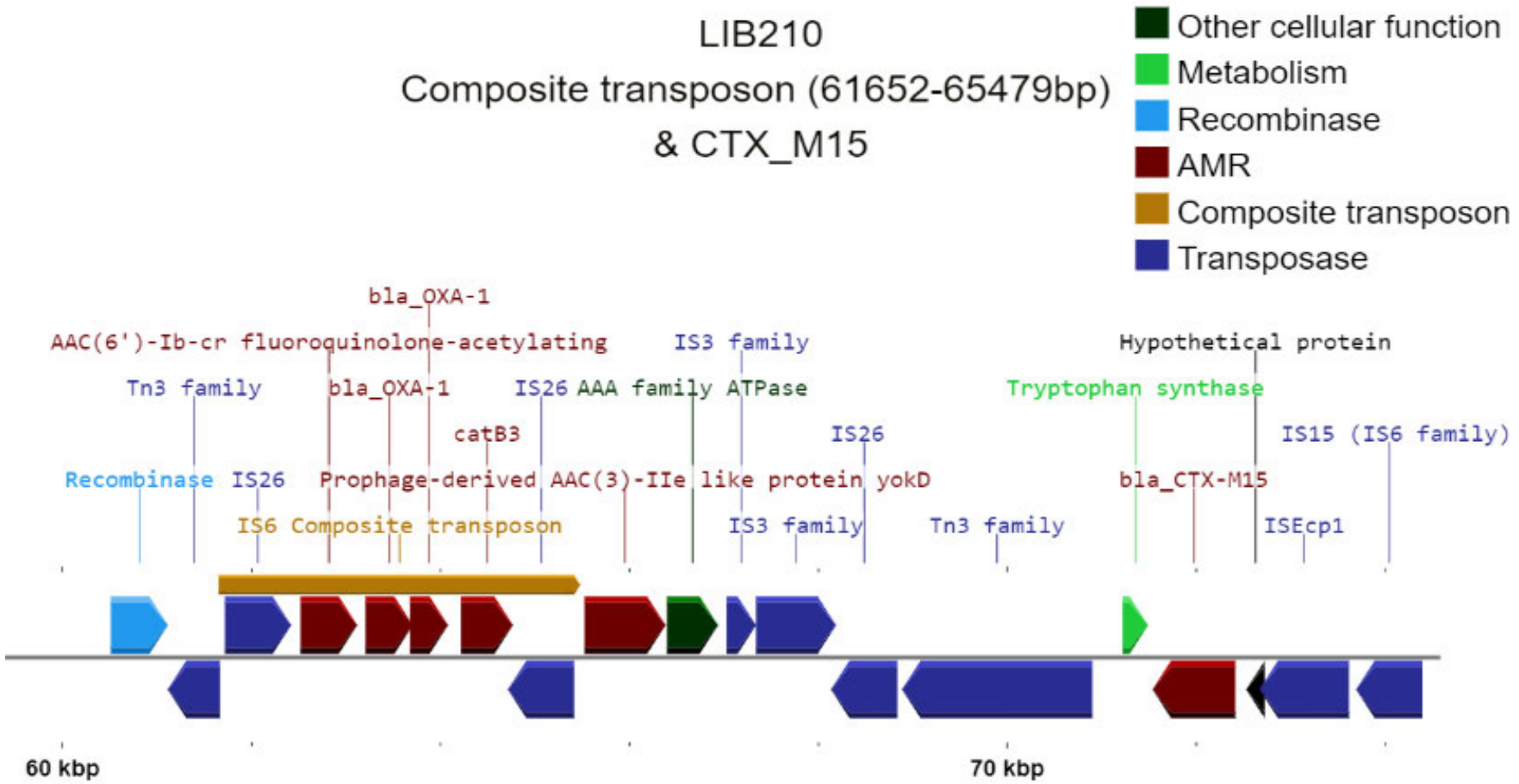


Figure 3.8. Linear map of a region of interest on the multiple drug resistant IncFII plasmid in *E. coli* isolate LIB210 identified using MGEFinder. Putative genes are coloured according to the predicted functions of their products. The yellow region features an IS26 composite transposon which harbours an *AAC(6')-Ib-cr* fluoroquinolone-acetylating gene, *bla*OXA-*1* genes and *catB3* genes. Immediately adjacent to this transposon is an *aac(3)-IIa* gene (red). Further upstream from this region and flanked by transposases is a *bla*CTX-*M15* gene (red). The direction of the arrows indicates the orientation of gene transcription.

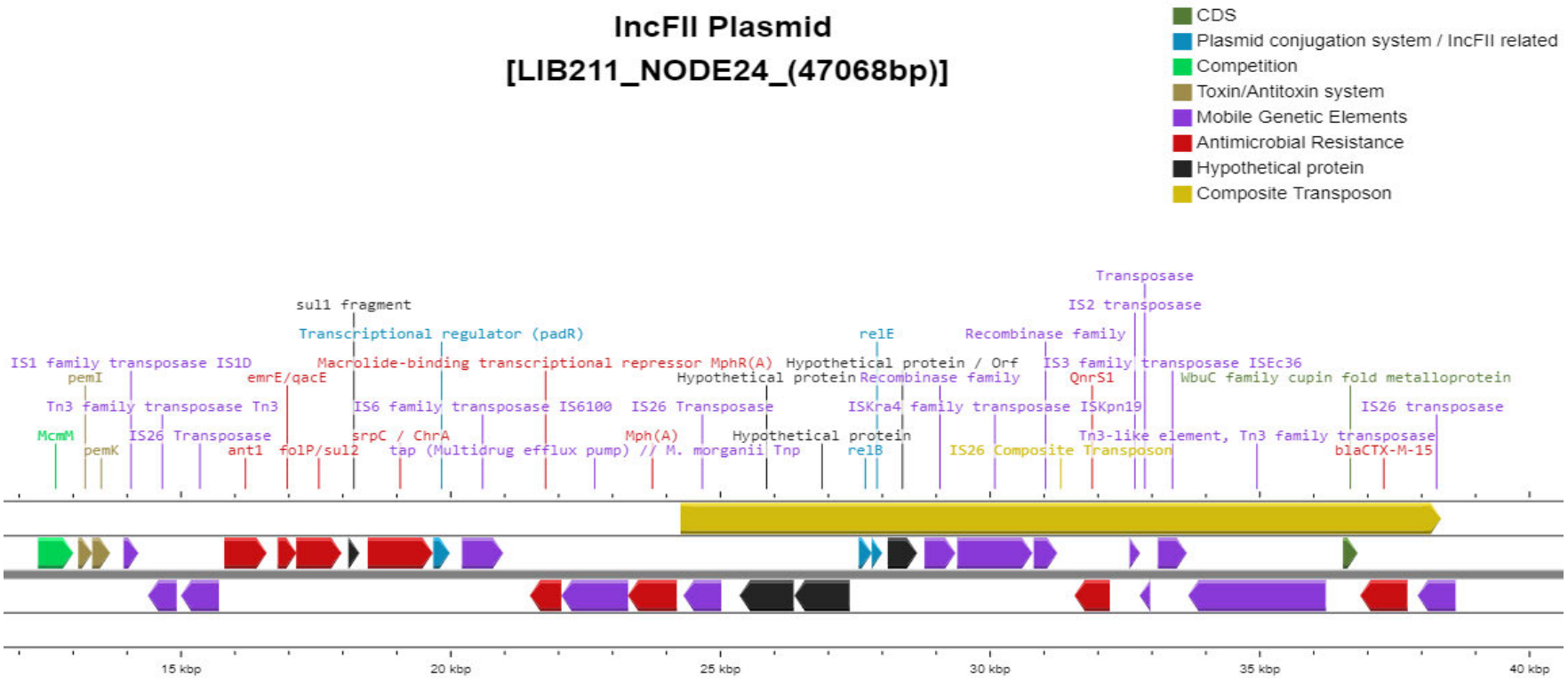


Figure 3.9. Linear map of a region of interest on the genome of *E. coli* isolate LIB211 that includes multiple antimicrobial resistance genes including adjacent genes *ant1*, *qacE*, *sul2* and *ChrA* (red) with Mph/R(A) (red) further upstream identified using MGEFinder on the CGE platform. Putative genes are coloured according to the predicted functions of their products. The yellow region is identified as an IS26 composite transposon which harbours both a *bla*CTX-*M15* and *qnrS* resistance genes. The direction of the arrows indicates the orientation of gene transcription.

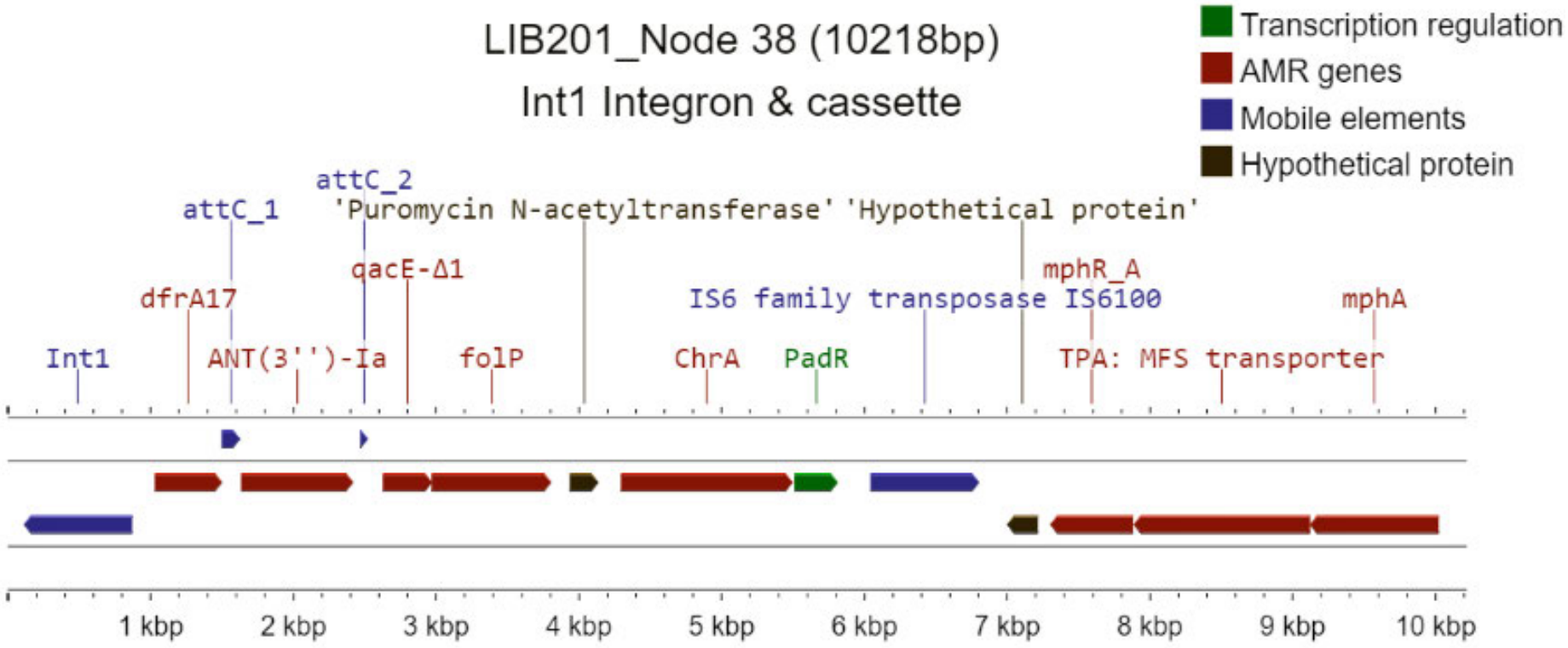


Figure 3.10. Linear map of a region of interest on the genome of *E. coli* isolate LIB201 that includes multiple antimicrobial resistance genes including *dfrA17*, *aad*A5, *qacE_delta1*, *folp*/*sul1* (red), genes associated with macrolide resistance including *mphA* and *mph(A)*, and integron/gene cassette system with integron-integrase gene *intI1* and two cassette recombination sites (*attC*), confirmed using IntegronFinder. Putative genes are coloured according to the predicted functions of their products. The direction of the arrows indicates the orientation of gene transcription.

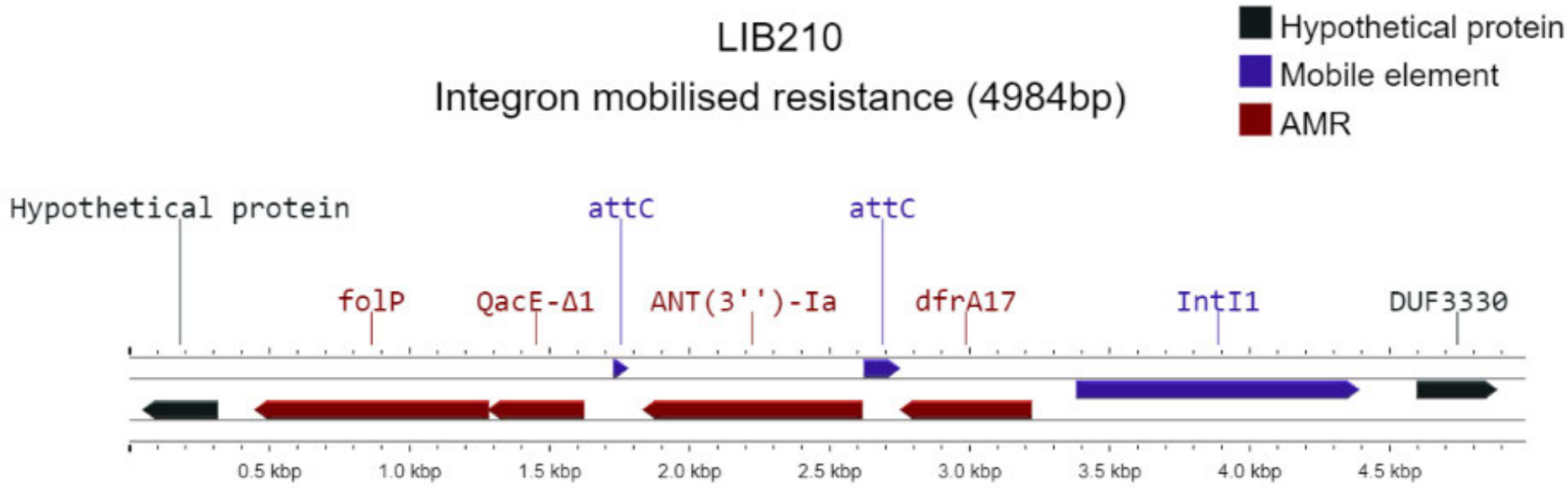


Figure 3.11. Linear map of a 4984bp region of interest on the genome of *E. coli* isolate LIB210 that includes a mobilised multidrug resistance region containing genes *folP/sul1*, *qacE_delta1*, *ANT(3'')-Ia* and *dfrA17* (red) flanked by an integron-integrase gene *int1* gene and two cassette recombination sites (*attC*) [confirmed using IntegronFinder}. Putative genes are coloured according to the predicted functions of their products. The direction of the arrows indicates the orientation of gene transcription.

4. Discussion

This study suggests ESBL production and identifies multidrug resistance in both human commensal and potentially pathogenic *E. coli* lineages as well as lineages with the potential to survive for extended periods in a hostile non-host environment, isolated from the treated wastewater effluent at East Calder Wastewater Treatment Plant.

4.1. **Coliform abundance in effluent and sediment.**

Overall, a difference could be observed in coliform population structure. The river sediment had a lower abundance of coliforms and had a significantly lower proportion of resistant *E. coli* and resistant Non-*E. coli* coliforms (NECC) compared to the effluent (Table 3.1). However, a delay in processing the sediment samples may explain the latter.

The abundance of *E. coli* determined for the WWTP effluent in the present study (2.0×10^2 cfu/1ml) was similar to that reported in other studies, for example Verburg *et al*. (2019) reported 1×10^2cfu/ml and Zanotto *et al*. (2016) reported 1.5×10^2cfu/ml. However, Ouatarra *et al*. (2011) reported a $\log_{10}$ higher numbers (5.4×10^3cfu/ml) as did Osinka *et al*. (2016) who reported 1.6×10^3cfu/ml. A similar study by Turolla *et al*. (2018) found *E. coli* abundance in effluent from three WWTP's to be slightly less than 10-fold lower (1-10cfu/ml) than the present study. A study by Watkinson *et al*., (2007) which partially focussed on effluent exposed to chlorination treatment, reported *E. coli* numbers 2-4-fold lower than the present study (0.02-2.0 cfu/ml).

The abundance of NECC (4.0×10^4) in the sediment is more than 10-fold lower than in the effluent (2.85×10^5). The total abundance of NECC in the effluent fit well within the range (1.9-5.5 log cfu/ml) demonstrated for untreated, treated, freshwater and tap water established by Rocha *et al*. (2018) and the range found by Servais *et al*. (2007) determined from the effluent of 12 WWTP's with varying treatments.

While there are no set standards in the UK for WWTP effluent, the concentration of *E. coli* and NECC discharged from the WWTP in the current study (2849cfu/ml) would be just below the standard limit for discharge set at 3000cfu/ml in Japan (Wako *et al*., 2012). However, the limit for *E. coli* discharge directly to an inland bathing water is 900cfu/100ml (based on a 90-percentile evaluation) (The Bathing Waters (Scotland) Regulation 2008). This is 10-fold lower than the number reported (8300cfu/ml) in the effluent at East Calder. However, according to

the WHO, the numbers found in the present study would still be suitable for crop and pasture irrigation using spray or sprinkler systems (Blumenthal *et al*., 2000).

Overall, the numbers of NECC in the present study, as in other studies, demonstrated that *E. coli* abundance in the effluent was well above that of non-point sources (Ouattara *et al*., 2011), rivers closer to anthropogenic influence (Petit *et al*., 2017) and agricultural waters (Khan *et al*., 2007). It is possible that the association of bacteria with particulate matter would form a readily settleable fraction which subsequently settles out of the water column into the sediment. This translates into an accumulation of organic matter at a higher proportion in the sediment when compared with the proportion of organic matter, as Total Suspended Solids, in the effluent (Brown *et al*., 2019a, 2019b).

Previous research has indicated that diversity in *E. coli* concentrations between effluent/water and sediment is related to season and hydrological factors that influence sediment resuspension events which promotes increased interaction and mixing between these compartments (Jang *et al.,* 2015). At the time of sampling there were no major hydrological events occurring from the previous week. However, the proximity of the grab sample collection point to that of sediment extraction point would indicate continuous mixing, settling and suspension events leading to constant turbidity, as indicated by the sandy texture of the sediment, devoid of settled organic matter, although the exact size of particulate was not measured.

The lower proportion of *E. coli* as a percentage of total coliforms (0.2%) in the sediment compared to the effluent (2.9%) may demonstrate the influence of effluent properties on sediment communities and visa-versa. A proportion of the *E. coli p*opulation in the sediment may also be present in a viable but non-culturable (VBNC) state as these results appear a little unusual (Robben *et al*., 2018). Garzio-Hadzick *et al*., 2010). Previous studies indicate that an increase in abundance of *E. coli* is generally found in sediment due to the longer die off periods and lower inactivation rates of *E. coli* in sediments. This may in turn be due to the hospitable (high nutrient / low light) environment (Pachepshy and Shelton, 2011). Given the nutrient and particulate matter content in the effluent (not measured) and with extended *E. coli* survival rates known to occur in sediment, results contrary to the present study may have been assumed (Haller *et al*., 2009; Malham *et al*., 2014). However, the low (not measured) winter temperature (Smith *et al*., 2019) and sediment properties, physiochemical interactions

(Hassard *et al*., 2016; 2017) and bacterial surface charge (Wyness *et al*., 2018), could influence the abundance and proliferation of bacteria in the sediment. When taking into consideration a delay in processing of sediment samples (72 hours), this is the more likely explanation for the lower than assumed *E. coli* abundance.

4.2 Phenotyping and genotyping of ESBL producing *E. coli*.

The initial screening for cefotaxime resistant isolates indicates that the wastewater treatment plant at East Calder was discharging potentially pathogenic *E. coli* into the river Almond at the time of sampling. Resistance to multiple antibiotics was observed to occur in clinically recognised bacteraemia and bacteriuric strains (Dale and Woodford, 2015), identified using culture based phenotypic and genotypic screens.

4.2.1 *E. coli* species confirmation

The efficacy of genotyping assays using *gad*A and *uid*A specific primers used to confirm *E. coli* species from the McConkey selective media culture-based screen was demonstrated by the identification of a presumed NECC isolate as *E. coli* and confirms this as a necessary step in a screening process. This isolate was later identified as a B2 phylogroup member and an identical REP type to another isolate. However, other isolates were removed from the selection process due to an initial misidentification as *E. coli*. This again confirmed the necessity of this additional confirmation assay which may have revealed cross contamination occurring during the colony picking process. No relationship between sampling point and identification status could be conferred (Table 3.2).

4.2.2 Characterisation of *E. coli* strains and phylogroups.

Four of the nine cefotaxime resistant isolates were identified as the same REP type (Figure 3.2). These four clonal isolates all recovered from the effluent, were subsequently identified as belonging to phylogroup B1. However, (data was not shown) an additional REP-PCR screen did identify a fifth clonal isolate which was cefotaxime sensitive from the sediment. These results demonstrate the utility of this assay as a method to study clonal diversity in environmental isolates as reviewed by Ishii and Sadowsky (2009) and to identify relationships between microbial populations in different matrices and seasonal variations (Jang *et al*., 2015).

Phylogroup B1 is suggested to maintain a generalist lifestyle strategy and is commonly identified in the environment (Pachepsky, 2011; Petit *et al*., 2017). The presence of this cefotaxime resistant strain suggests it is dominant throughout the latter stages of wastewater treatment process with potential to persist in the receiving environment. The antibiogram of for isolate (Table 3.3) determined no other resistance than cefotaxime and may suggest a lower fitness cost associated with less resistant gene carriage and a possible strategy for environmental persistence (Johnson *et al*., 2003) which could in turn explain the higher abundance compared to the other isolates.

Two of the remaining isolates were also identified as clonally related, and, as mentioned in section 4.2.1 were both belonged to B2 phylogroup and were found in both the effluent and sediment. With the remaining three isolates all being different REP types and subsequently different phylogroups (C,D and F) the results suggest that while *E. coli* diversity is present in the effluent a single cefotaxime resistant clonal group, considered a human commensal and generalist (Picard *et al*., 1998; Gonzalez-Alba *et al*., 2019) is dominant. Any further assumptions are less than tentative due to the small sample size.

4.2.3 Evaluation and comparison of antimicrobial resistance and lineage.

The antibiogram screen identified resistance against multiple classes of antibiotics spread across all isolates except for the B1 clone. While all cefotaxime resistant isolates, confirmed by the antibiogram, carried *bla*CTX-M genes as confirmed by the PCR, additional β-lactam resistance was identified in the three bacteraemic strains including *bla*OXA in LIB210 (phylogroup C) and *bla*TEM in both LIB205 (phylogroup B2) and LIB211 (phylogroup D). Multiple Drug Resistance was identified in the antibiogram but was restricted to those phylogroups more frequently associated with bacteraemia (Clermont *et al*., 2016).

Although LIB201 and LIB205 are identified as clonal with identical antibiogram resistance profiles, this was not reflected in ESBL gene carriage, as only LIB205 carried a *bla*TEM gene. Further, the isolates differed in *bla*CTX gene carriage with LIB201 identified with a group CTX-M-9 gene and LIB205 with a group CTX-M-1 gene. It was this difference in CTX-M group which influenced the selection for WGS.

Both phenotyping and genotyping identified an assortment of phylogroups with pathogenic potential as retaining a higher abundance of ARGs. Phylogroups B2, C and D are well

renowned for causing bacteraemia associated with AMR (Clermont *et al*., 2016). Two isolates that were identified as belonging to this phylogroup, LIB201 and LIB211, were identified as carrying multiple resistance. Both these isolates also maintained ciprofloxacin and cefotaxime resistance suggesting that these were indeed isolates of clinical concern (Dale and Woodford, 2015). With LIB201 and LIB205 being phylogroup B2 members and LIB211 being a phylogroup D member, further genotyping of the sequence type (ST) determined by MLST was performed but failed to identify them as clinically recognised pathogenic clonal lineages.

To maintain diversity, the phylogroup F isolate (LIB213) while having an identical result to the antibiogram screen as the phylogroup D isolate LIB211, was not identified with additional ESBL resistance, this isolate was not investigated with WGS (Table 3.3). Additionally, only one of the phylogroup B2 isolate clones, LIB201 was carried forward due to the difference in CTX-M group. This meant that only four isolates, LIB201, LIB202, LIB210 and LIB211 were selected for WGS analysis.

4.3 In-silico analysis.

The selected four isolates with their genomes fully sequenced, were subjected to *in silico* analysis.

4.3.1 Identification of ST lineages.

All four isolates were identified as belonging to potentially pathogenic ST groups (Table 3.5). Two isolates (LIB201 and LIB211) were identified as belonging to renowned pathogenic clonal lineages ST131 and ST69 (Riley 2014), respectively. Both these ST groups were screened for using PCR but were not detected and only identified as such using MLST finder. This indicated that while the positive control in the PCR assay worked correctly, an unidentified error occurred during processing such as processing of incorrect or non-viable controls. Due to time constraints this was not repeated.

The remaining two resistant isolates were also identified as potentially pathogenic lineages, ST58 and ST44 (a sub-lineage of ST10), both commonly identified in bacteraemia. However, these two ST groups were not screened for during genotyping. Isolate LIB202 (ST58) was the only resistant isolate identified as a human commensal and generalist although has pathogenic potential.

4.3.2 Analysis of individual isolate ST lineages and pathotype.

4.3.2.1 Isolate LIB201.

In comparison to the other 4 resistant isolates, LIB201 (ST131), can be classified as a higher risk ExPEC member and is renowned for global distribution with strong associations to CTX-M ESBL genes (Nicolas-Chanoine *et al*., 2014; Zakour *et al*., 2016). The O25:H4 serotype of this isolate is concerning as in the UK ST131 with a O25 serotype is commonly associated with both urinary tract and bloodstream infections (Ciesielcuk *et al*., 2016).

Previous analysis identified a CTX-M9 group gene carried by LIB201 and with additional analysis with ResFinder this was shown to be a *bla*CTX-M-27 gene. Subsequently, LIB201 fits the profile of a newly emerged H30/R1 sub-clade, C1-M27, according to Matsumura *et al*. (2016). To further confirm this assumption, a BLAST identified 100% homology with prophage 1 region that provides designation to the C1-M27 clade (Matsumura *et al*., 2016). The classification was further confirmed by the class C virotype (Blanco *et al*., 2013).

What makes identification of this isolate especially interesting is that not only is this a newly emerging clone, but CTX-M27 (ST131 H30R1/Clade 1) variants are more commonly identified in East Asia whereas in Europe CTX-M15 (ST131 H30Rx/Clade 2) (Matsumura *et al*., 2015; Merino *et al*., 2018). This observation may suggest that the other B2 isolate identified during the preliminary screen (LIB205) may in fact be the C2 variant due to the differing CTX-M group enzyme.

There appears to be a higher frequency of the C1-sub-group across host species and environmental compartments (Jamborova *et al*., 2018). However, in general, this C1 sub-group is not as frequently identified in clinical samples when compared to the C2 sub-group, again a little surprising as the class C virotype assigned to this strain is suggestive of a metabolic profile directed towards infection over colonisation (Blanco *et al*., 2010) and studies have identified that the presence of both papA and KPSIIM virulence genes are positively correlated with fatal mortality rates in murine model studies (Merino *et al*., 2020).

4.3.2.2 Isolate LIB211.

Similar to LIB201, isolate LIB211 is also identified as a similar higher risk ExPEC member and globally distributed (Riley, 2014). The sequence type assigned to LIB211 (ST69), like ST131 and is considered as one of 4 priority MLST's (although the list is expanding) associated with EXPEC and UPEC in the UK and is commonly identified in Europe (Roer *et al*., 2017; Johnson *et al*.,

2019). However, ST69 is regarded as ecologically distinct from the other MLST's and is often commonly identified circulating within the community and hospital associated infections (Goswami *et al*., 2018; Fox *et al*., 2020). The serotype of O15:H18 has been pathotyped to Enteroaggregative *E. coli* (EAEC) in a study by Beutin *et al*. (2004) and an identical serotype and CH type are recognised in another clinical study although with a differing resistance pattern (Marchetti *et al*., 2020). More recently a similar isolate was found circulating in Wild Ibis in Australia (Wyrsch *et al*., 2019) with similar pan-genome features such as plasmid type and Insertion sequence family.

4.3.2.3 Isolate LIB210.

The sequence type of isolate LIB210 (ST44) is a clonal complex member of ST10 (STc10). STc10 is more often associated with phylogroup A but is also commonly identified in phylogroup B1 (Racicot-Bergeron *et al*., 2012; Zingali *et al*., 2020; Zhao *et al*., 2022) and generally there is relative stability with STc phylogroup distribution. However, LIB210 is a phylogroup C member and some of the classification schemes may rely on the older Clermont typing which would not distinguish phylogroup A from C (Clermont *et al*., 2013).

While STc10 is subject to antigenic diversity, the O101:H4 fimH54 with ciprofloxacin resistance and the ESBL encoding genes in LIB210 present a unique diversification pattern compared to other members of this ST10 complex. The O101:H4 serotype is known for causing an outbreak of haemorrhagic diarrhoea in 2011 in Germany (Bielaszewska *et al*., 2011). However, this was not caused by a STc10 strain but a rare ST678 isolate with a Shiga toxin encoding prophage, which appears absent in LIB210.

In the UK, between 2013 and 2014, during a comparison of ESBL producing *E. coli* isolates causing clinical bacteraemia and ESBL producing *E. coli* from different non-clinical settings (where ST10 isolates dominated cattle and slurry isolates) STc10 accounted for 3 out of 253 cases of bacteraemia (in contrast ST131 accounted for 188/253). This again also highly specific on serotype. Generally, STc10 is considered antimicrobial susceptible and a benign intestinal commensal (Manges and Johnson, 2012). However, there have been instances of MDR with some ESBL production in this group with the O101 serotype isolated from Human faecal swabs and Animal cloacal swabs and the subsequent retail products (Racicot-Bergeron *et al*., 2012; Zhao *et al*., 2022). A similar ST44 (cc10) isolate was detected in cockroaches identified to carry CTX-M group 1 enzymes and a TEM enzyme and is suggestive of presenting a household

conduit for this strain to enter humans (Obeng-Nkrumah *et al*., 2019). Subsequently, STc10 is also emerging as a potentially pathogenic clone (Manges and Johnson, 2012).

4.3.2.4. Isolate LIB202.

While commonly regarded as benign commensals, virulence gene acquisition has also been identified in members of the B1 phylogroup (Johnson and Stell, 2000; Jang *et al*., 2013; Poolman, 2017) which may warrant concern if these commensals have acquired a more pathogenic genotype (Wright *et al*., 2008; Gillings *et al*., 2008; An *et al*., 2018). Isolate LIB202 was identified as ST58, which resides within the ST155 clonal complex and comprises solely of B1 phylogroup members and shares similar phylogroup conservation in this respect to LIB201, the ST131 isolate. ST58 is frequently identified with ESBL carriage possibly due to the presence of ESBL carrying plasmids, as demonstrated by this study, and is recognised, globally, as a sequence type common in bacteraemia (McKinnon *et al*., 2018; Mamani *et al*., 2019). It has also been isolated from food-producing animals as well as wild and captive animals (Ahmed *et al*., 2017; Wyrsch *et al*., 2020; Rojas-Jiminez *et al*., 2021). Interestingly, a study by Ahmed *et al*. (2017) identified potentially mobilised ESBL production (*bla*TEM-1b genes) in ST58 isolates in the absence of antibiotic selective pressure.

However, the serogroup is currently unknown and could not be identified using Enterobase but may indicate a potentially non-pathogenic clade (https://enterobase.warwick.ac.uk/species/ index/ecoli). Although lack of abundance of virulence factors is common for this phylogroup, the virulence genes that were present were interesting as they may facilitate the pathogenicity of this isolate given that they are often associated with UPEC and ExPEC pathotypes (Clermont *et al*., 2017).

4.3.3 Characterization of Antimicrobial resistance.

4.3.3.1 ESBL resistance in LIB202 mobilised by an IncI1 plasmid.

When challenged with the antibiogram, isolate LIB202 (B1) showed resistance to 3 classes of B-lactams antibiotics only, Ampicillin, Cefotaxime and Aztreonam a commonly prescribed monobactam. *In silico* analysis using ResFinder confirmed the AST Antibiogram and the genotyping results, identifying a *bla*CTX-M1 gene as responsible for all three confirmed resistance phenotypes.

The *bla*CTX-M1 gene was mobilised by an Inc1 plasmid and was the only resistance mobilising plasmid type other than IncF types identified. This was the only plasmid identified in this isolate and subsequently another reason this isolate is particularly interesting. This plasmid family is common in Enterobacteriaceae and is frequently associated with the carriage of ESBL's in *E. coli* (Day *et al*., 2016), and is considered clinically relevant (Caratolli *et al*., 2018; Petrin *et al*., 2021). Antibiotic selection pressure from both clinical and veterinary practice has resulted in these plasmids developing into major vehicles of ESBL dissemination (Carattoli *et al*., 2018). A previous study by Sallem *et al*. (2013) identified this plasmid type in ST58 isolates circulating in humans and companion animals and, similarly to the LIB202 isolate, these were also carrying CTX-M1 group enzymes.

The *bla*CTX-M1 gene was identified adjacent to an ISEcp1 transposition unit in amongst the IncI1 conjugation machinery which is a common occurrence in these plasmids (Zong *et al*., 2015; Stosic *et al*., 2021) as well as both narrow and broad host range plasmids (Canton and Coque, 2006). However, in LIB202, the integration of the *bla*CTX-M-1 gene into the ISEcp1 transposition unit appeared incomplete when compared to these previous studies. However, the position of both the *bla*CTX-M1 gene and the transposition unit within the plasmid conjugative machinery, the shufflon region, does appear highly similar with these studies (Zong *et al*., 2015; Stosic *et al*., 2021).

Flanking the *bla*CTX-M1 gene was a tryptophan synthetase gene which may ameliorate the cost of carriage (Dunn *et al*., 2008). Indole acts as an interspecies signal molecule known to modulate processes such as AMR, tolerance to acids as well as virulence (Zarkan *et al*., 2020). More interestingly it has been shown to inhibit conjugation so potentially the tryptophan synthase gene may affect both *bla*CTX-M gene expression and ameliorate the cost associated with carriage and conjugation of this plasmid (Zarkan *et al*., 2020; Xiong *et al*., 2021). This may be an interesting target to conduct knock-out studies and abundance studies.

4.3.3.2 Multiple drug resistance is mobilised by IncF type plasmids pathogenic lineages.

The three remaining isolates (LIB201, LIB210 and LIB211) were identified with multiple plasmids. This observation, together with the multiple drug resistance profile, provided a clear contrast between these potentially pathogenic lineages and the more general commensal LIB202. The ESBL's were mobilised in IncF type plasmids except for LIB202 (Table 3.7).

The resistant profiles of the three IncF type carrying isolates suggests that this multiple drug resistance commonly shuffled across the species as similar regions of resistance were observed between isolates. IncF plasmids are frequently identified in Enterobacteriaceae with carriage of MDR also commonplace (Carattoli, 2009) however, through BLASTN analysis these remain partial sequences thereby limiting interpretation.

However, due to the sequencing method which used short read only, replicons were recognised as separate contigs, and full resolution was not possible. This may be due to extended regions of repetitive nucleotide sequences greater that 250bp in length and is a limitation of this study method. However, characterisation was possible with the coverage provided as some of the replicon types were identified in contigs containing resistance gene regions.

The IncFII replicon associated with MDR in LIB201 shared homology with a plasmid (pRSB107) isolated from an uncultured bacterium that also demonstrated MDR. Interestingly, this plasmid was also isolated from a WWTP (Szczepanowski *et al*., 2005) suggesting its persistence in this environment. The other replicons (IncFIA and IncFIB) identified in LIB201 as separate contigs, could be resolved to be a single plasmid. If this was the case then these replicons could in fact be the same RepFIA, FIB and FII plasmid backbone identified by (Szczepanowski *et al*., 2005) and indicate a typical sized (>100kb) IncF type plasmid (Villa *et al*., 2009).

4.3.3.3 IS6 family insertion elements mobilise drug resistance in the IncF type plasmids.

The IS6 family insertion sequences are commonly associated with the spread and rearrangement of antimicrobial resistance genes (Razavi *et al*., 2020), and form intimate relationships with plasmids (Harmer and Hall, 2019; Varani *et al*., 2021). Multiple IS6 family insertion sequences were identified in all three and were the predominant mobile elements associated with resistance in these isolates. The abundance of these family of insertion elements confirms the findings by Razavi *et al*. (2020) who also identified their high abundance in similar wastewater environments.

IS6 family insertion sequence derived composite transposons mobilised ESBLs in all three of the isolates. In LIB210 and LIB211 these transposons contained both ESBL and quinolone resistance. Although the ESBL resistance in LIB210 was associated with both the *bla*OXA-1 and

*bla*CTX-M15 genes only the *bla*OXA gene was carried in the transposon. Of concern here is a potential growing trend of Carbapenemase resistance in these strains with both plasmid mediated quinolone resistance (*aac*(6")-lb-cr) and through mutations (Phan *et al*., 2022). However, although the ST131 isolate had mutation mediated quinolone resistance only the ST69 isolate carried both plasmid and mutation mediated quinolone resistance. However, in LIB 210

The *bla*CTX-M15 gene in LIB210 was adjacent to the transposon (Figure 3.7). Similar to LIB202 this *bla*CTX-M gene was flanked by a tryptophan synthase gene and also was embedded in an ISEcp1 transposition unit, and both were adjacent to an additional transposon (Tn3 family) flanked by IS6 family insertion sequence types. Although this was not identified as a composite transposon it may suggest that the insertion into LIB202 is a truncated insertion element which has become potentially fixed in the shufflon region although did not share complete homology with the study by Stosic *et al*. (2021) However, it is indicative of similar resistance mobility occurring between these two plasmid types. Additionally, it suggests a relationship between indole scavenging and CTX-M gene maintenance (Xiong *et al*., 2021).

A recurring resistance pattern was observed as a region of resistance genes identified as carried by a Class 1 integron in isolates LIB201 and LIB210 (section 4.3.3.4). These same genes were present in LIB211 in the same order forming a MDR region but mobilised by a different transposase in a region flanking the composite transposon (Figure 3.8). This same drug resistant region appears conserved as it appears commonly in IncFII plasmids found in ST131 strains with the same association with IS6 family insertion sequence (Johnson *et al*., 2016).

4.3.3.4 LIB201 and LIB210 carry an identical multidrug resistant Class1 integron.

The integron identified on isolate LIB201 (Figure 3.9) shared 100% homology with similar Class 1 integrons isolated from multiple similar ST131 C1/H30R isolates from both humans and pigs carrying the *bla*CTX-M27 gene in Japan (Matsuo *et al*., 2020) and unpublished studies in Israel and the UK. An identical integron was also detected in isolate LIB210 (Figure 3.10) and carried the same four resistance genes: sul1, *qac*E_delta1, ANT(3")-Ia and *dfr*A17 (Figures 3.9 and 3.10). The identification of a highly conserved fused region between attC site and *qac*EΔ this integron is suggested by studies (Quintela-Baluja *et al*., 2021) to be indicative of and responsive to anthropogenic activity (Gillings *et al*., 2014; Lehmann *et al*., 2016) typifying its identification in this study. The presence in both animals and humans in the above studies

highlights a concern that this is maintained by clinically recognised pathogens with global distribution in the non-clinical environment (Quintela-Baluja *et al*., 2021).

This integron was featured as a small contig for isolates LIB201 and LIB210 but may be mobilised on the IncF type plasmids as reported in other studies and subsequently readily circulating in other reciprocal bacterial species (Villa *et al*., 2009; Carattoli, 2009). This study agrees with other studies that suggest the Class 1 integrons are suitable markers for anthropogenic pollution (Gillings, 2017; Quintela-Baluja *et al*., 2021).

5. Conclusion and future studies.

This study identified potentially pathogenic *E. coli* with mobilised multidrug resistance in the treated effluent being discharged into the river Almond from the wastewater treatment plant at East Calder. The isolation study showed that approximately 2.4% of the total *E. coli* recovered had ESBL activity. A confirmation assay using *E. coli* specific primers identified additional isolates as *E. coli* indicating the use of both coliform selective and McConkey selective media should be followed up with genotype screening. Genes relevant to ESBL production were initially identified in multiple phylogroups, but multidrug resistance as identified by the Kirby-Bauer assay occurred in the more pathogenic phylogroups. An additional isolate was observed with a single *bla*CTX-M1 gene that accounted for approximately 50% of all the resistant isolates although this occurred in a commensal strain. This gene was mobilised by an IncI1 plasmid.

Resistance to multiple antibiotics including ESBL, quinolones, macrolides, aminoglycosides, disinfectants, and folate pathway agonists was observed in IncF type plasmids carried by clinically recognised bacteraemia and bacteriuric strains of *E. coli*. Three globally renowned ST were identified ST131 and ST69 as well as ST44 (which resides in clonal complex ST10). The ST131 isolate was identified as the global pandemic clone (ST131 H30R1/Clade 1) and demonstrates the continuing spread of this clonal lineage. Resistance genes were commonly associated with IS6 family transposable elements and similar patterns could be observed in different lineages and was suggestive of interspecies transmission. Additionally, two Class1 integrons with identical gene cassettes were identified in the ST131 and ST69 lineages.

Although a limited number of isolates were assessed, their presence is clinically concerning and warrants further investigation of WWTP's and receiving surface waters due to high transmissibility of ESBL and other resistance genes.

Future characterisation studies.

As the B1 clone is observed to be the dominant species to persist in the outflow and Phylogroup C (LIB210) is considered a sister lineage, if not indistinguishable from Phylogroup B1, a future study may investigate the genetic environments of these two isolates. This would include traits that can be linked to non-host survival and potential cross over between the two. Certainly, virulence factor acquisition such as those pertaining to biofilm formation would indicate a strategy evolved towards non-host survival. A further key area would be plasmid identification, resistance carriage by these plasmids and incompatibility between plasmids and their host phylogroups. Another study direction could be to further characterise the plasmids using both short and long read sequencing platforms to further resolve the IncF plasmids. In addition to this, the study would also benefit from conjugation assays to determine the transmissibility of the resistance genes identified. Further knock-out studies could also assess whether the tryptophan synthase gene, identified in two isolates, lowers the cost of *bla*CTX-M gene carriage. Other biological characterisation of the wastewater could also be advantageous such as chemical analysis as there is the potential legacy effect from previous industrial activity within the vicinity that may drive some of these selection processes.

Being that CTX-M27 (ST131 H30R1/Clade 1) variants are more commonly identified in East Asia, a future study may be directed towards confirming this observation to determine whether this clonal groups is in fact circulating within the representative population and being released into the environment.

Mucilage populations were isolated but not quantified. This plant associated matrix has also received little research. One study suggested that while diversity may decrease in this matrix, mucilage may provide a specific environmental niche, rich in nutrients, which may support or encourage diversity (Quero *et al*., 2015). This non-hostile environment with potentially high frequency shedding events may in turn facilitate survival of *E. coli* and persistence further downstream subsequently suggesting further research is required of this particular matrix.

References

Ageitos, J. M., Sánchez-Pérez, A., Calo-Mata, P., & Villa, T. G. (2017). Antimicrobial peptides (AMPs): Ancient compounds that represent novel weapons in the fight against bacteria. *Biochemical Pharmacology*, 133, 117–138). https://doi.org/10.1016/j.bcp.2016.09.018

Ahmed, S., Olsen, J. E., & Herrero-Fresno, A. (2017). The genetic diversity of commensal *Escherichia coli* strains isolated from nonantimicrobial treated pigs varies according to age group. *PLoS ONE*, 12. https://doi.org/10.1371/journal.pone.0178623

Alcock, B. P., Raphenya, A. R., Lau, T. T. Y., Tsang, K. K., Bouchard, M., Edalatmand, A., Huynh, W., Nguyen, A. 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 Research, 48(D1), D517-D525

Amos, G. C. A., Hawkey, P. M., Gaze, W. H. & Wellington, E. M. (2014). Waste water effluent contributes to the dissemination of CTX-M-15 in the natural environment. *Journal of Antimicrobial Chemotherapy*, 69, 1785–1791. https://doi.org/10.1093/jac/dku079

Amos, G. C. A., Ploumakis, S., Zhang, L., Hawkey, P. M., Gaze, W. H. & Wellington, E. M. H. (2018). The widespread dissemination of integrons throughout bacterial communities in a riverine system. *ISME Journal*, 12, 681–691. https://doi.org/10.1038/s41396-017-0030-8

Amos, G. C., Gozzard, E., Carter, C. E., Mead, A., Bowes, M. J., Hawkey, P. M., Zhang, L., Singer, A. C., Gaze, W. H. & Wellington, E. M. H. (2015). Validated predictive modelling of the environmental resistome. *ISME Journal*, 9, 1467–1476. https://doi.org/10.1038/ismej.2014.237

Auguet, O., Pijuan, M., Borrego, C. M., Rodriguez-Mozaz, S., Triadó-Margarit, X., Giustina, S. V. della, & Gutierrez, O. (2017). Sewers as potential reservoirs of antibiotic resistance. *Science*

of the Total Environment, *605–606*, 1047–1054. https://doi.org/10.1016/j.scitotenv.2017.06.153

ben Zakour, N. L., Alsheikh-Hussain, A. S., Ashcroft, M. M., Khanh Nhu, N. T., Roberts, L. W., Stanton-Cook, M., Schembri, M. A. & Beatsona, S. A. (2016). Sequential acquisition of virulence and fluoroquinolone resistance has shaped the evolution of *Escherichia coli* ST131. *MBio*, 7. https://doi.org/10.1128/mBio.00347-16

Bengtsson-Palme, J., Kristiansson, E. & Larsson, D. G. J. (2018). Environmental factors influencing the development and spread of antibiotic resistance. *FEMS Microbiology Reviews*, 42, 68–80. https://doi.org/10.1093/femsre/fux053

Berche, P. (2012). Louis Pasteur, from crystals of life to vaccination. *Clinical Microbiology and Infection* 18, 1–6. https://doi.org/10.1111/j.1469-0691.2012.03945.x

Bergeron, C. R., Prussing, C., Boerlin, P., Daignault, D., Dutil, L., Reid-Smith, R. J., Zhanel, G. G. & Manges, A. R. (2012). Chicken as reservoir for extraintestinal pathogenic *Escherichia coli* in Humans, Canada. *Emerging Infectious Diseases*, 18, 415–421. https://doi.org/10.3201/eid1803.111099

Beutin, L., Tao, J., Feng, L., Krause, G., Zimmermann, S., Gleier, K., Xia, Q. & Wang, L. (2005). Sequence analysis of the Escherichia coli O15 antigen gene cluster and development of a PCR assay for rapid detection of intestinal and extraintestinal pathogenic *E. coli* O15 strains. *Journal of Clinical Microbiology*, 43, 703–710. https://doi.org/10.1128/JCM.43.2.703-710.2005

Bevan, E. R., Jones, A. M. & Hawkey, P. M. (2017). Global epidemiology of CTX-M β-lactamases: Temporal and geographical shifts in genotype. *Journal of Antimicrobial Chemotherapy*, 72, 2145–2155. https://doi.org/10.1093/jac/dkx146

Bielaszewska, M., Mellmann, A., Zhang, W., Köck, R., Bauwens, A., Karch, H., Fruth, A., Bauwens, A. & Peters, G. (2011). Characterisation of the *Escherichia coli* strain associated with an outbreak of haemolytic uraemic syndrome in Germany, 2011: a microbiological study. *Lancet Infect Dis*, 11, 671–676. https://doi.org/10.1016/S1473

Blount, Z. D. (2015). The unexhausted potential of *E. coli*. *ELife*, 4. https://doi.org/10.7554/eLife.05826

Bonten, M., Johnson, J. R., van den Biggelaar, A. H. J., Georgalis, L., Geurtsen, J., de Palacios, P. I., Gravenstein, S., Verstraeten, T., Hermans, P. & Poolman, J. T. (2021). Epidemiology of *Escherichia coli* Bacteremia: A Systematic Literature Review. *Clinical Infectious Diseases*, 72, 1211–1219. Oxford University Press. https://doi.org/10.1093/cid/ciaa210

Bürgmann, H., Frigon, D., Gaze, W. H., Manaia, C. M., Pruden, A., Singer, A. C., Smets, B. F. & Zhang, T. (2018). Water and sanitation: An essential battlefront in the war on antimicrobial resistance. *FEMS Microbiology Ecology*, 94. https://doi.org/10.1093/femsec/fiy101

Burke, L., Humphreys, H. & Fitzgerald-Hughes, D. (2012). The revolving door between hospital and community: extended-spectrum beta-lactamase-producing *Escherichia coli* in Dublin. *Journal of Hospital Infection*, 81, 192-198.

Bush, K. (2018). Past and present perspectives on β-lactamases. *Antimicrobial Agents and Chemotherapy*, 62. *American Society for Microbiology*. https://doi.org/10.1128/AAC.01076-18

Cantón, R., Akóva, M., Carmeli, Y., Giske, C. G., Glupczynski, Y., Gniadkowski, M., Livermore, D. M., Miriagou, V., Naas, T., Rossolini, G. M., Samuelsen, Seifert, H., Woodford, N., Nordmann, P., Poirel, L., Bogaerts, P., Navon-Venezia, S. & Cornaglia, G. (2012). Rapid evolution and spread of carbapenemases among *Enterobacteriaceae* in Europe. *Clinical Microbiology and Infection*, 18, 413–431. https://doi.org/10.1111/j.1469-0691.2012.03821.x

Cantón, R., Novais, A., Valverde, A., Machado, E., Peixe, L., Baquero, F. & Coque, T. M. (2008). Prevalence and spread of extended-spectrum β-lactamase-producing *Enterobacteriaceae* in Europe. *Clinical Microbiology and Infection*, 14, 144–153. https://doi.org/10.1111/j.1469-0691.2007.01850.x

Carattoli, A. (2009). Resistance plasmid families in *Enterobacteriaceae*. *Antimicrobial Agents and Chemotherapy*, 53, 2227–2238. https://doi.org/10.1128/AAC.01707-08

Carattoli, A. (2013). Plasmids and the spread of resistance. *International Journal of Medical Microbiology*, 303, 298–304. https://doi.org/10.1016/J.IJMM.2013.02.001

Carattoli, A., Villa, L., Fortini, D. & García-Fernández, A. (2021). Contemporary IncI1 plasmids involved in the transmission and spread of antimicrobial resistance in Enterobacteriaceae. *Plasmid* 118. https://doi.org/10.1016/j.plasmid.2018.12.001

Che, Y., Xia, Y., Liu, L., Li, A. D., Yang, Y. & Zhang, T. (2019). Mobile antibiotic resistome in wastewater treatment plants revealed by Nanopore metagenomic sequencing. *Microbiome*, 7. https://doi.org/10.1186/s40168-019-0663-0

Ciesielczuk, H., Jenkins, C., Chattaway, M., Doumith, M., Hope, R., Woodford, N. & Wareham, D. W. (2016). Trends in ExPEC serogroups in the UK and their significance. *European Journal of Clinical Microbiology and Infectious Diseases*, 35, 1661–1666. https://doi.org/10.1007/s10096-016-2707-8

Clermont, O., Couffignal, C., Blanco, J., Mentré, F., Picard, B. & Denamur, E. (2017). Two levels of specialization in bacteraemic *Escherichia coli* strains revealed by their comparison with commensal strains. *Epidemiology and Infection*, 145, 872–882. https://doi.org/10.1017/S0950268816003010

Conte, D., Palmeiro, J. K., da Silva Nogueira, K., de Lima, T. M. R., Cardoso, M. A., Pontarolo, R., Degaut Pontes, F. L. & Dalla-Costa, L. M. (2017). Characterization of CTX-M enzymes, quinolone resistance determinants, and antimicrobial residues from hospital sewage, wastewater treatment plant, and river water. *Ecotoxicology and Environmental Safety*, 136, 62–69. https://doi.org/10.1016/j.ecoenv.2016.10.031

Corno, G., Yang, Y., Eckert, E. M., Fontaneto, D., Fiorentino, A., Galafassi, S., Zhang, T. & di Cesare, A. (2019). Effluents of wastewater treatment plants promote the rapid stabilization of the antibiotic resistome in receiving freshwater bodies. *Water Research*, 158, 72–81. https://doi.org/10.1016/j.watres.2019.04.031

Czekalski, N., Gascón Díez, E. & Bürgmann, H. (2014). Wastewater as a point source of antibiotic-resistance genes in the sediment of a freshwater lake. *ISME Journal*, 8, 1381–1390. https://doi.org/10.1038/ismej.2014.8

Czekalski, N., Sigdel, R., Birtel, J., Matthews, B. & Bürgmann, H. (2015). Does human activity impact the natural antibiotic resistance background? Abundance of antibiotic resistance genes in 21 Swiss lakes. *Environment International*, 81, 45–55. https://doi.org/10.1016/j.envint.2015.04.005

D'Andrea, M. M., Arena, F., Pallecchi, L. & Rossolini, G. M. (2013). CTX-M-type β-lactamases: A successful story of antibiotic resistance. *International Journal of Medical Microbiology*, 303, 305–317. https://doi.org/10.1016/j.ijmm.2013.02.008

Dale, A. P. & Woodford, N. (2015). Extra-intestinal pathogenic *Escherichia coli* (ExPEC): Disease, carriage and clones. *Journal of Infection*, 71, 615–626. https://doi.org/10.1016/j.jinf.2015.09.009

Davies, J. & Davies, D. (2010). Origins and Evolution of Antibiotic Resistance. *Microbiology and Molecular Biology Reviews*, 74, 417-433. doi : 10.1128/MMBR.00016-10

de Toro, M., Fernández, J., García, V., Mora, A., Blanco, J., de La Cruz, F. & Rodicio, M. R. (2017). Whole genome sequencing, molecular typing and in vivo virulence of OXA-48-producing *Escherichia coli* isolates including ST131 H30-Rx, H22 and H41 subclones. *Scientific Reports*, 7. https://doi.org/10.1038/s41598-017-12015-0

Dealtry, S., Ding, G. C., Weichelt, V., Dunon, V., Schlüter, A., Martini, M. C., del Papa, M. F., Lagares, A., Amos, G. C. A., Wellington, E. M. H., Gaze, W. H., Sipkema, D., Sjöling, S., Springael, D., Heuer, H., van Elsas, J. D., Thomas, C. & Smalla, K. (2014). Cultivation-independent screening revealed hot spots of IncP-1, IncP-7 and IncP-9 plasmid occurrence in different environmental habitats. *PLoS ONE*, 9. https://doi.org/10.1371/journal.pone.0089922

Diene, S. M., Pinault, L., Armstrong, N., Azza, S., Keshri, V., Khelaifia, S., Chabrière, E., Caetano-Anolles, G., Rolain, J. M., Pontarotti, P. & Raoult, D. (2020). Dual rnase and β-lactamase activity of a single enzyme encoded in archaea. *Life*, 10. https://doi.org/10.3390/life10110280

Dorado-García, A., Smid, J. H., van Pelt, W., Bonten, M. J. M., Fluit, A. C., van den Bunt, G., Wagenaar, J. A., Hordijk, J., Dierikx, C. M., Veldman, K. T., de Koeijer, A., Dohmen, W., Schmitt, H., Liakopoulos, A., Pacholewicz, E., Lam, T. J. G. M., Velthuis, A. G., Heuvelink, A., Gonggrijp,

M. A., van Duijkeren, E., van Hoek, A. H. A. M., de Roda Husman, A. M., Blaak, H., Havelaar, A. H., Mevius, D. J., Heederik, D. J. J. (2018). Molecular relatedness of ESBL/AmpC-producing *Escherichia coli* from humans, animals, food and the environment: A pooled analysis. *Journal of Antimicrobial Chemotherapy*, 73, 339–347.doi : 10.1093/jac/dkx397

Drawz, S. M. & Bonomo, R. A. (2010). Three decades of β-lactamase inhibitors. *Clinical Microbiology Reviews*, 23, 160–201. https://doi.org/10.1128/CMR.00037-09

Drieux, L., Haenn, S., Moulin, L. & Jarlier, V. (2016). Quantitative evaluation of extendedspectrum β-lactamase-producing *Escherichia coli* strains in the wastewater of a French teaching hospital and relation to patient strain. *Antimicrobial Resistance and Infection Control*, 5. https://doi.org/10.1186/s13756-016-0108-5

Dunn, M. F., Niks, D., Ngo, H., Barends, T. R. M. & Schlichting, I. (2008). Tryptophan synthase: the workings of a channeling nanomachine. *Trends in Biochemical Sciences* 33, 254–264. https://doi.org/10.1016/j.tibs.2008.04.008

Ellaby, N., Doumith, M., Hopkins, K. L., Woodford, N. & Ellington, M. J. (2019). Emergence of diversity in carbapenemase-producing *Escherichia coli* ST131, England, January 2014 to June 2016. *Eurosurveillance*, 24. https://doi.org/10.2807/1560-7917.ES.2019.24.37.1800627

Fleming, A. (1929) On the Antibacterial Action of Cultures of a Penicillium, with Special Reference to Their Use in the Isolation of *B. Influenzae*. *British Journal of Experimental Pathology,* 10, 226.

Fox, S., Goswami, C., Holden, M., Connolly, J. P. R., Mordue, J., O'Boyle, N., Roe, A., Connor, M., Leanord, A. & Evans, T. J. (2020). A highly conserved complete accessory *Escherichia coli* type III secretion system 2 is widespread in bloodstream isolates of the ST69 lineage. *Scientific Reports*, 10. https://doi.org/10.1038/s41598-020-61026-x

Friday, J. (1974). A Microscopic incident in a monumental struggle: Huxley and antibiosis in 1875. *The British Journal for the History of Science*, 7, 61–71. https://doi.org/10.1017/S0007087400012863

Gao, H., Zhang, L., Lu, Z., He, C., Li, Q. & Na, G. (2018). Complex migration of antibiotic resistance in natural aquatic environments. *Environmental Pollution*, 232, 1–9. https://doi.org/10.1016/j.envpol.2017.08.078

Geurtsen, J., de Been, M., Weerdenburg, E., Zomer, A., McNally, A. & Poolman, J. (2022). Genomics and pathotypes of the many faces of *Escherichia coli*. *FEMS Microbiology Reviews*, 46. https://doi.org/10.1093/femsre/fuac031

Gillings, M. R. (2013). Evolutionary consequences of antibiotic use for the resistome, mobilome, and microbial pangenome. *Frontiers in Microbiology*, 4. https://doi.org/10.3389/fmicb.2013.00004

Gillings, M. R. (2014). Integrons: Past, Present, and Future. *Microbiology and Molecular Biology Reviews*, 78, 257–277. https://doi.org/10.1128/mmbr.00056-13

Gillings, M. R. (2017). Class 1 integrons as invasive species. *Current Opinion in Microbiology* (Vol. 38, pp. 10–15). https://doi.org/10.1016/j.mib.2017.03.002

Gillings, M. R. (2018). DNA as a Pollutant: the Clinical Class 1 Integron. *Current Pollution Reports*, 4, 49–55. https://doi.org/10.1007/s40726-018-0076-x

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 Journal*, 9, 1269–1279. https://doi.org/10.1038/ismej.2014.226

World Health Organization. (2023). Global antimicrobial resistance and use surveillance system (GLASS) report: 2022 Available at: *https://www.who.int/publications-detail-redirect/9789240062702.*

Gonzalez-Alba, J. M., Baquero, F., Cantón, R. & Galán, J. C. (2019). Stratified reconstruction of ancestral *Escherichia coli* diversification. *BMC Genomics*, 20. https://doi.org/10.1186/s12864-019-6346-1

Goswami, C., Fox, S., Holden, M., Connor, M., Leanord, A. & Evans, T. J. (2018). Genetic analysis of invasive *Escherichia coli* in Scotland reveals determinants of healthcare-associated

versus community acquired infections. *Microbial Genomics*, 4. https://doi.org/10.1099/mgen.0.000190

Grant, J. R. & Stothard, P. (2008). The CGView Server: a comparative genomics tool for circular genomes. *Nucleic Acids Research*, 36. https://doi.org/10.1093/nar/gkn179

Griffin, P. M., Manges, A. R. & Johnson, J. R. (2012). Food-borne origins of *Escherichia coli* causing extraintestinal infections. *Clinical Infectious Diseases*, 55, 712–719. https://doi.org/10.1093/cid/cis502

Hall, J. P. J., Brockhurst, M. A. & Harrison, E. (2017). Sampling the mobile gene pool: Innovation via horizontal gene transfer in bacteria. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372. https://doi.org/10.1098/rstb.2016.0424

Hamed, R. B., Gomez-Castellanos, J. R., Henry, L., Ducho, C., McDonough, M. A. & Schofield, C. J. (2013). The enzymes of β-lactam biosynthesis. *Natural Product Reports*, 30, 21–107. https://doi.org/10.1039/c2np20065a

Harmer, C. J. & Hall, R. M. (2019). An analysis of the IS6/IS26 family of insertion sequences: Is it a single family? *Microbial Genomics*, 5. https://doi.org/10.1099/mgen.0.000291

Hassard, F., Andrews, A., Jones, D. L., Parsons, L., Jones, V., Cox, B. A., Daldorph, P., Brett, H., McDonald, J. E. & Malham, S. K. (2017). Physicochemical factors influence the abundance and culturability of human enteric pathogens and fecal indicator organisms in estuarine water and sediment. *Frontiers in Microbiology*, 8. https://doi.org/10.3389/fmicb.2017.01996

Hassard, F., Gwyther, C. L., Farkas, K., Andrews, A., Jones, V., Cox, B., Brett, H., Jones, D. L., McDonald, J. E. & Malham, S. K. (2016). Abundance and distribution of enteric bacteria and viruses in coastal and estuarine sediments-A review. *Frontiers in Microbiology*. 7. https://doi.org/10.3389/fmicb.2016.01692

Hassoun-Kheir, N., Stabholz, Y., Kreft, J. U., de la Cruz, R., Dechesne, A., Smets, B. F., Romalde, J. L., Lema, A., Balboa, S., García-Riestra, C., Torres-Sangiao, E., Neuberger, A., Graham, D., Quintela-Baluja, M., Stekel, D. J., Graham, J., Pruden, A., Nesme, J., Sørensen, S. J., Hought R., Paul, M. (2021). EMBRACE-WATERS statement: Recommendations for reporting of studies on

antimicrobial resistance in wastewater and related aquatic environments. *One Health*, 13. https://doi.org/10.1016/j.onehlt.2021.100339

Hawkey, P. M. (2015). Multidrug-resistant Gram-negative bacteria: A product of globalization. *Journal of Hospital Infection* 89, 241–247. https://doi.org/10.1016/j.jhin.2015.01.008

Health Protection Scotland An Official Statistics publication for Scotland Health Protection Scotland. (n.d.). https://www.statisticsauthority.gov.uk/osr/code-of-practice/

Hopkins, S. (2010). English surveillance programme for antimicrobial utilisation and resistance Antibiotic Guardian [View project: Optimising antibiotic prescribing]. https://doi.org/10.13140/RG.2.1.1149.8004

Huijbers, P. M. C., Flach, C. F. & Larsson, D. G. J. (2019). A conceptual framework for the environmental surveillance of antibiotics and antibiotic resistance. *Environment International*, 130. https://doi.org/10.1016/j.envint.2019.05.074

Hultman, J., Tamminen, M., Pärnänen, K., Cairns, J., Karkman, A. & Virta, M. (2018). Host range of antibiotic resistance genes in wastewater treatment plant influent and effluent. *FEMS Microbiology Ecology*, 94. https://doi.org/10.1093/FEMSEC/FIY038

Humeniuk, C., Arlet, G., Gautier, V., Grimont, P., Labia, R. & Philippon, A. (2002). β-lactamases of *Kluyvera ascorbata*, probable progenitors of some plasmid-encoded CTX-M types. *Antimicrobial Agents and Chemotherapy*, 46, 3045–3049. https://doi.org/10.1128/AAC.46.9.3045-3049.2002

Hutinel, M., Huijbers, P. M. C., Fick, J., Åhrén, C., Larsson, D. G. J. & Flach, C. F. (2019). Population-level surveillance of antibiotic resistance in *Escherichia coli* through sewage analysis. *Eurosurveillance*, 24. https://doi.org/10.2807/1560-7917.ES.2019.24.37.1800497

Ishii, S. & Sadowsky, M. J. (2009). Applications of the rep-PCR DNA fingerprinting technique to study microbial diversity, ecology and evolution: Minireview. *Environmental Microbiology*, 11, 733–740. https://doi.org/10.1111/j.1462-2920.2008.01856.x

Jang, J., Di, D. Y. W., Han, D., Unno, T., Lee, J. H., Sadowsky, M. J. & Hur, H. G. (2015). Dynamic changes in the population structure of *Escherichia coli* in the Yeongsan River basin of South Korea. *FEMS Microbiology Ecology*, 91. https://doi.org/10.1093/femsec/fiv127

Jang, J., Suh, Y. S., Di, D. Y. W., Unno, T., Sadowsky, M. J. & Hur, H. G. (2013). Pathogenic *Escherichia coli* strains producing extended-spectrum β-lactamases in the Yeongsan river basin of South Korea. *Environmental Science and Technology*, 47, 1128–1136. https://doi.org/10.1021/es303577u

Johnson, J. R., Clermont, O., Johnston, B., Clabots, C., Tchesnokova, V., Sokurenko, E., Junka, A. F., Maczynska, B. & Denamur, E. (2014). Rapid and specific detection, molecular epidemiology, and experimental virulence of the O16 subgroup within *Escherichia coli* sequence type 131. *Journal of Clinical Microbiology*, 52, 1358–1365. https://doi.org/10.1128/JCM.03502-13

Johnson, J. R., Johnston, B. D., Porter, S. B., Clabots, C., Bender, T. L., Thuras, P., Trott, D. J., Cobbold, R., Mollinger, J., Ferrieri, P., Drawz, S. & Banerjee, R. (2019). Rapid emergence, subsidence, and molecular detection of *Escherichia coli* sequence type 1193-fimH64, a new disseminated multidrug-resistant commensal and extraintestinal pathogen. *Journal of Clinical Microbiology*, 57. https://doi.org/10.1128/JCM.01664-18

Johnson, J. R., Johnston, B. D., Porter, S., Thuras, P., Aziz, M. & Price, L. B. (2019). Accessory traits and phylogenetic background predict *Escherichia coli* extraintestinal virulence better than does ecological source. *Journal of Infectious Diseases*, 219, 121–132. https://doi.org/10.1093/infdis/jiy459

Johnson, J. R., Johnston, B., Thuras, P., Launer, B., Sokurenko, E. v. & Miller, L. G. (2016). *Escherichia coli* Sequence Type 131 H 30 Is the Main Driver of Emerging Extended-Spectrum-β-Lactamase-Producing *E. coli* at a Tertiary Care Center . *MSphere*, 1. https://doi.org/10.1128/msphere.00314-16

Johnson, J. R., Tchesnokova, V., Johnston, B., Clabots, C., Roberts, P. L., Billig, M., Riddell, K., Rogers, P., Qin, X., Butler-Wu, S., Price, L. B., Aziz, M., Nicolas-Chanoine, M. H., Debroy, C., Robicsek, A., Hansen, G., Urban, C., Platell, J., Trott, D. J., Zhanel, G., Weissman, S. J., Cookson,

B. T., Fang, F. C., Limaye, A. P., Scholes, D., Chattopadhyay, S., Hooper, D. C. & Sokurenko, E. V. (2013). Abrupt emergence of a single dominant multidrug-resistant strain of *Escherichia coli*. *Journal of Infectious Diseases*, 207, 919-928. https://doi.org/10.1093/infdis/jis933

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 H 30 Subclone of *Escherichia coli* Sequence Type 131 . *MSphere*, 1. https://doi.org/10.1128/msphere.00121-16

Johnson, T. J., Elnekave, E., Miller, E. A., Munoz-Aguayo, J., Figueroa, C. F., Johnston, B., Nielson, D. W., Logue, C. M. & Johnson, J. R. (2019). Phylogenomic Analysis of Extraintestinal Pathogenic *Escherichia coli* Sequence Type 1193, an Emerging Multidrug-Resistant Clonal Group. *Antimicrobial Agents and Chemotherapy*, 63. https://doi.org/10.1128/AAC.01913-18

Jutkina, J., Marathe, N. P., Flach, C. F. & Larsson, D. G. J. (2018). Antibiotics and common antibacterial biocides stimulate horizontal transfer of resistance at low concentrations. *Science of the Total Environment*, *616–617*, 172–178. https://doi.org/10.1016/j.scitotenv.2017.10.312

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 Research*, 27, 1437–1449. https://doi.org/10.1101/gr.216606.116

Kaspar, C. W., Hartman, P. A. & Benson, A. K. (1987). Coagglutination and Enzyme Capture Tests for Detection of *Escherichia coli* 1-Galactosidase, 3-Glucuronidase, and Glutamate Decarboxylase. *Applied and Environmental Microbiology* 53. 10.1128/aem.53.5.1073-1077.1987

Larsson, D. G. J. & Flach, C. F. (2022). Antibiotic resistance in the environment. *Nature Reviews Microbiology*. 20, 257–269. https://doi.org/10.1038/s41579-021-00649-x

Lau, S. H., Kaufmann, M. E., Livermore, D. M., Woodford, N., Willshaw, G. A., Cheasty, T., Stamper, K., Reddy, S., Cheesbrough, J., Bolton, F. J., Fox, A. J. & Upton, M. (2008). UK

epidemic *Escherichia coli* strains A-E, with CTX-M-15 β-lactamase, all belong to the international O25:H4-ST131 clone. *Journal of Antimicrobial Chemotherapy*, 62, 1241–1244. https://doi.org/10.1093/jac/dkn380

Laxminarayan, R. (2022). The overlooked pandemic of antimicrobial resistance. *The Lancet* 399, 606–607. https://doi.org/10.1016/S0140-6736(22)00087-3

Lehmann, K., Bell, T., Bowes, M. J., Amos, G. C. A., Gaze, W. H., Wellington, E. M. H. & Singer, A. C. (2016). Trace levels of sewage effluent are sufficient to increase class 1 integron prevalence in freshwater biofilms without changing the core community. *Water Research*, 106, 163–170. https://doi.org/10.1016/j.watres.2016.09.035

Lekunberri, I., Balcázar, J. L. & Borrego, C. M. (2018). Metagenomic exploration reveals a marked change in the river resistome and mobilome after treated wastewater discharges. *Environmental Pollution*, 234, 538–542. https://doi.org/10.1016/j.envpol.2017.12.001

Leonard, A. F. C., Zhang, L., Balfour, A. J., Garside, R. & Gaze, W. H. (2015). Human recreational exposure to antibiotic resistant bacteria in coastal bathing waters. *Environment International*, 82, 92–100. https://doi.org/10.1016/j.envint.2015.02.013

Leonard, A. F. C., Zhang, L., Balfour, A. J., Garside, R., Hawkey, P. M., Murray, A. K., Ukoumunne, O. C. & Gaze, W. H. (2018). Exposure to and colonisation by antibiotic-resistant *E. coli* in UK coastal water users: Environmental surveillance, exposure assessment, and epidemiological study (Beach Bum Survey). *Environment International*, 114, 326–333. https://doi.org/10.1016/j.envint.2017.11.003

Li, J., Cheng, W., Xu, L., Strong, P. J. & Chen, H. (2015). Antibiotic-resistant genes and antibiotic-resistant bacteria in the effluent of urban residential areas, hospitals, and a municipal wastewater treatment plant system. *Environmental Science and Pollution Research*, 22, 4587–4596. https://doi.org/10.1007/s11356-014-3665-2

Linares, J. F., Gustafsson, I., Baquero, F. & Martinez, J. L. (2006). Antibiotics as intermicrobial signaling agents instead of weapons. *Proceedings of the National Academy of Sciences of the United States of America,* 103, 19484-19489. https://doi.org/10.1073/pnas.0608949103

Long, G. S., Klunk, J., Duggan, A. T., Tapson, M., Giuffra, V., Gazzè, L., Fornaciari, A., Duchene, S., Fornaciari, G., Clermont, O., Denamur, E., Golding, G. B. & Poinar, H. (2022). A 16th century *Escherichia coli* draft genome associated with an opportunistic bile infection. *Communications Biology*, 5. https://doi.org/10.1038/s42003-022-03527-1

Low, M., Neuberger, A., Hooton, T. M., Green, M. S., Raz, R., Balicer, R. D. & Almog, R. (2019). Association between urinary community-acquired fluoroquinolone-resistant *Escherichia coli* and neighbourhood antibiotic consumption: a population-based case-control study. *The Lancet Infectious Diseases*, 19, 419–428. https://doi.org/10.1016/S1473-3099(18)30676-5

Lu, J., Wang, Y., Li, J., Mao, L., Nguyen, S. H., Duarte, T., Coin, L., Bond, P., Yuan, Z. & Guo, J. (2018). Triclosan at environmentally relevant concentrations promotes horizontal transfer of multidrug resistance genes within and across bacterial genera. *Environment International*, 121. 1217–1226. https://doi.org/10.1016/j.envint.2018.10.040

Marano, R. B. M., Fernandes, T., Manaia, C. M., Nunes, O., Morrison, D., Berendonk, T. U., Kreuzinger, N., Tenson, T., Corno, G., Fatta-Kassinos, D., Merlin, C., Topp, E., Jurkevitch, E., Henn, L., Scott, A., Heß, S., Slipko, K., Laht, M., Kisand, V., Di Cesare, A., Karaolia, P., Michael, S. G., Petre, A. L., Rosal, R., Pruden, A., Riquelme, V., Agüera, A., Esteban, B., Luczkiewicz, A., Kalinowska, A., Leonard, A., Gaze, W. H., Adegoke, A. A., Stenstrom, T. A., Pollice, A., Salerno, C., Schwermer, C. U., Krzeminski, P., Guilloteau, H., Donner, E., Drigo, B., Libralato, G., Guida, M., Bürgmann, H., Beck, K., Garelick, H., Tacão, M., Henriques, I., Martínez-Alcalá, I., Guillén-Navarro, J. M., Popowska, M., Piotrowska, M., Quintela-Baluja, M., Bunce, J. T., Polo-López, M. I., Nahim-Granados, S., Pons, M. N., Milakovic, M., Udikovic-Kolic, N., Ory, J., Ousmane, T., Caballero, P., Oliver, A., Rodriguez-Mozaz, S., Balcazar, J. L., Jäger, T., Schwartz, T., Yang, Y., Zou, S., Lee, Y., Yoon, Y., Herzog, B., Mayrhofer, H., Prakash, O., Nimonkar, Y., Heath, E., Baraniak, A., Abreu-Silva, J., Choudhury, M., Munoz, L. P., Krizanovic, S., Brunetti, G., Maile-Moskowitz, A., Brown, C. & Cytryn, E. (2020). A global multinational survey of cefotaxime-resistant coliforms in urban wastewater treatment plants. *Environmental International,* 144, 106035. https://doi.org/10.1016/j.envint.2020.106035

Mathers, A. J., Peirano, G. & Pitout, J. D. D. (2015). The role of epidemic resistance plasmids and international high- risk clones in the spread of multidrug-resistant *Enterobacteriaceae*. *Clinical Microbiology Reviews*, 28, 565–591. https://doi.org/10.1128/CMR.00116-14

Matsumura, Y., Johnson, J. R., Yamamoto, M., Nagao, M., Tanaka, M., Takakura, S. & Ichiyama, S. (2014). CTX-M-27- and CTX-M-14-producing, ciprofloxacin-resistant *Escherichia coli* of the H30 subclonal group within ST131 drive a Japanese regional ESBL epidemic. *Journal of Antimicrobial Chemotherapy*, 70, 1639–1649. https://doi.org/10.1093/jac/dkv017

Matsumura, Y., Pitout, J. D. D., Gomi, R., Matsuda, T., Noguchi, T., Yamamoto, M., Peirano, G., de Vinney, R., Bradford, P. A., Motyl, M. R., Tanaka, M., Nagao, M., Takakura, S. & Ichiyama, S. (2016). Global *Escherichia coli* sequence type 131 clade with *bla*CTX-M-27 gene. *Emerging Infectious Diseases*, 22, 1900–1907. https://doi.org/10.3201/eid2211.160519

Matsuo, N., Nonogaki, R., Hayashi, M., Wachino, J. I., Suzuki, M., Arakawa, Y. & Kawamura, K. (2020). Characterization of blaCTX-M-27/F1:A2:B20 Plasmids Harbored by *Escherichia coli* Sequence Type 131 Sublineage C1/H30R Isolates Spreading among Elderly Japanese in Nonacute-Care Settings. *Antimicrobial Agents and Chemotherapy,* 64*,* e00202-20. https://doi.org/10.1128/AAC.00202-20

McKinnon, J., Roy Chowdhury, P. & Djordjevic, S. P. (2018). Genomic analysis of multidrug-resistant *Escherichia coli* ST58 causing urosepsis. *International Journal of Antimicrobial Agents*, 52, 430–435. https://doi.org/10.1016/j.ijantimicag.2018.06.017

Medeiros, A. A. (1997). β-lactamases: Quality and resistance. *Clinical Microbiology and Infection*, *3*(SUPPL. 4). https://doi.org/10.1016/s1198-743x(14)65030-8

Morris, D., Boyle, F., Buckley, V., Xu, L., Hanahoe, B., Hawkey, P. & Cormican, M. (2009). CTX-M enzymes are the predominant extended-spectrum β-lactamases produced by *Enterobacteriaceae* in Ireland. *Journal of Antimicrobial Chemotherapy*, 64, 864–866. https://doi.org/10.1093/jac/dkp297

Murray, C. J. L., 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., Achalapong,

S., Agarwal, R., Akech, S., Albertson, S., Amuasi, J., Andrews, J., Aravkin, A., Ashley, E., Babin, F.-X., Bailey, F., Baker, S., Basnyat, B., Bekker, A., Bender, R., Berkley, J. A., Bethou, A., Bielicki, J., Boonkasidecha, S., Bukosia, J., Carvalheiro, C., Castañeda-Orjuela, C., Chansamouth, V., Chaurasia, S., Chiurchiù, S., Chowdhury, F., Donatien, R. C., Cook, A. J., Cooper, B., Cressey, T. R., Criollo-Mora, E., Cunningham, M., Darboe, S., Day, N. P. J., De Luca, M., Dokova, K., Dramowski, A., Dunachie, S. J., Bich, T. D., Eckmanns, T., Eibach, D., Emami, A., Feasey, N., Fisher-Pearson, N., Forrest, K., Garcia, C., Garrett, D., Gastmeier, P., Giref, A. Z., Greer, R. C., Gupta, V., Haller, S., Haselbeck, A., Hay, S. I., Holm, M., Hopkins, S., Hsia, Y., Iregbu, K. C., Jacobs, J., Jarovsky, D., Javanmardi, F., Jenney, A. W. J., Khorana, M., Khusuwan, S., Kissoon, N., Kobeissi, E., Kostyanev, T., Krapp, F., Krumkamp, R., Kumar, A., Kyu, H. H., Lim, C., Lim, K., Limmathurotsakul, D., Loftus, M. J., Lunn, M., Ma, J., Manoharan, A., Marks, F., May, J., Mayxay, M., Mturi, N., Munera-Huertas, T., Musicha, P., Mussi-Pinhata, M. M., Naidu, R. N., Nakamura, T., Nanavati, R., Nangia, S., Newton, P., Ngoun, C., Novotney, A., Nwakanma, D., Obiero, C. W., Ochoa, T. J., Olivas-Martinez, A., Olliaro, P., Ooko, E., Ortiz-Brizuela, E., P, A. Y., Perrone, C., Phe, T., Phommasone, K., Plakkal, N., Ponce-de-Leon, A., Raad, M., Ramdin, T., Rattanavong, S., Riddell, A., Roberts, T., Robotham, J. V., Roca, A., Rosenthal, V. D., Rudd, K. E., Russell, N., Sader, H. S., Saengchan, W., Schnall, J., Scott, J. A. G., Seekaew, S., Sharland, M., Shivamallappa, M., Sifuentes-Osornio, J., Simpson, A. J., Steenkeste, N., Stewardson, A. J., Stoeva, T., Tasak, N., Thaiprakong, A., Thwaites, G., Tigoi, C., Turner, C., Turner, P., van Doorn, H. R., Velaphi, S., Vongpradith, A., Vongsouvath, M., Vu, H., Walsh, T., Walson, J. L., Waner, S., Wangrangsimakul, T., Wannapinij, P., Wozniak, T., Young Sharma, T. E. M., Yu, K. C., 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. *The Lancet*, 399, 629–655. https://doi.org/10.1016/S0140-6736(21)02724-0

Nicolas-Chanoine, M. H., Bertrand, X. & Madec, J. Y. (2014). *Escherichia coli* ST131, an intriguing clonal group. *Clinical Microbiology Reviews*, 27, 543–574. https://doi.org/10.1128/CMR.00125-13

Nordmann, P., Naas, T. & Poirel, L. (2011). Global spread of carbapenemase producing *Enterobacteriaceae*. *Emerging Infectious Diseases*, 17, 1791–1798. https://doi.org/10.3201/eid1710.110655

Numberger, D., Ganzert, L., Zoccarato, L., Mühldorfer, K., Sauer, S., Grossart, H. P. & Greenwood, A. D. (2019). Characterization of bacterial communities in wastewater with enhanced taxonomic resolution by full-length 16S rRNA sequencing. *Scientific Reports*, 9. https://doi.org/10.1038/s41598-019-46015-z

O'Flaherty, E., Solimini, A., Pantanella, F. & Cummins, E. (2019). The potential human exposure to antibiotic resistant *Escherichia coli* through recreational water. *Science of the Total Environment*, 650, 786–795. https://doi.org/10.1016/j.scitotenv.2018.09.018

Obeng-Nkrumah, N., Labi, A. K., Blankson, H., Awuah-Mensah, G., Oduro-Mensah, D., Anum, J., Teye, J., Kwashie, S. D., Bako, E., Ayeh-Kumi, P. F. & Asmah, R. (2019). Household cockroaches carry CTX-M-15-, OXA-48- And NDM-1-producing *enterobacteria* and share beta-lactam resistance determinants with humans. *BMC Microbiology*, 19. https://doi.org/10.1186/s12866-019-1629-x

Ouattara, N. K., Passerat, J. & Servais, P. (2011). Faecal contamination of water and sediment in the rivers of the Scheldt drainage network. *Environmental Monitoring and Assessment*, 183, 243–257. https://doi.org/10.1007/s10661-011-1918-9

Pal, C., Bengtsson-Palme, J., Kristiansson, E. & Larsson, D. G. J. (2015). Co-occurrence of resistance genes to antibiotics, biocides and metals reveals novel insights into their co-selection potential. *BMC Genomics*, 16. https://doi.org/10.1186/s12864-015-2153-5

Pärnänen, K. M. M., Narciso-da-Rocha, C., Kneis, D., Berendonk, T. U., Cacace, D., Do, T. T., Elpers, C., Fatta-Kassinos, D., Henriques, I., Jaeger, T., Karkman, A., Martinez, J. L., Michael, S. G., Michael-Kordatou, I., O'Sullivan, K., Rodriguez-Mozaz, S., Schwartz, T., Sheng, H., Sørum, H., Stedtfeld, R. D., Tiedje, J. M., Giustina, S. V. D., Walsh, F., Vaz-Moreira, I., Virta, M. & Manaia, C. M. (2019). Antibiotic resistance in European wastewater treatment plants mirrors the pattern of clinical antibiotic resistance prevalence. *Science Advances,* 5, eaau9124. https://doi.org/10.1126/sciadv.aau9124

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-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. *Antimicrobial Agents and Chemotherapy*, 63. https://doi.org/10.1128/AAC.00823-19

Perry, J., Waglechner, N. & Wright, G. (2016). The prehistory of antibiotic resistance. *Cold Spring Harbor Perspectives in Medicine*, *6*. https://doi.org/10.1101/cshperspect.a025197

Petty, N. K., Zakour, N. L. B., Stanton-Cook, M., Skippington, E., Totsika, M., Forde, B. M., Phan, M. D., Moriel, D. G., Peters, K. M., Davies, M., Rogers, B. A., Dougan, G., Rodriguez-Baño, J., Pascual, A., Pitout, J. D. D., Upton, M., Paterson, D. L., Walsh, T. R., Schembri, M. A. & Beatson, S. A. (2014). Global dissemination of a multidrug resistant *Escherichia coli* clone. *Proceedings of the National Academy of Sciences of the United States of America*, 111, 5694–5699. https://doi.org/10.1073/pnas.1322678111

Phan, M.-D., Peters, K. M., Alvarez Fraga, L., Wallis, S. C., Hancock, S. J., Nhu, N. T. K., Forde, B. M., Bauer, M. J., Paterson, D. L., Beatson, S. A., Lipman, J. & Schembri, M. A. (2022). Plasmid-mediated ciprofloxacin resistance imparts a selective advantage on *Escherichia coli* ST131. *Antimicrobial Agents and Chemotherapy*, 66, e0214621. https://doi.org/10.1128/AAC.02146-21

Picard, B., Garcia, J. S., Gouriou, S., Duriez, P., Brahimi, N., Bingen, E., Elion, J. & Denamur, E. (1999). The link between phylogeny and virulence in *Escherichia coli* extraintestinal infection? *Infection and Immunity*, 67, 546–553. https://doi.org/10.1128/iai.67.2.546-553.1999

Pitout, J. D. D. (2012). Extraintestinal pathogenic *Escherichia coli:* A combination of virulence with antibiotic resistance. *Frontiers in Microbiology* 3. https://doi.org/10.3389/fmicb.2012.00009

Polo, D., Quintela-Baluja, M., Corbishley, A., Jones, D. L., Singer, A. C., Graham, D. W. & Romalde, J. L. (2020). Making waves: Wastewater-based epidemiology for COVID-19 – approaches and challenges for surveillance and prediction. *Water Research*, 186. https://doi.org/10.1016/j.watres.2020.116404

Poolman, J. T. (2016). *Escherichia coli*. *International Encyclopedia of Public Health* (pp. 585–593). Elsevier. https://doi.org/10.1016/B978-0-12-803678-5.00504-X

Prieto Riquelme, M. V., Garner, E., Gupta, S., Metch, J., Zhu, N., Blair, M. F., Arango-Argoty, G., Maile-Moskowitz, A., Li, A. D., Flach, C. F., Aga, D. S., Nambi, I. M., Larsson, D. G. J., Bürgmann, H., Zhang, T., Pruden, A. & Vikesland, P. J. (2022). Demonstrating a Comprehensive Wastewater-Based Surveillance Approach That Differentiates Globally Sourced Resistomes. *Environmental Science and Technology*, 56, 14982–14993. https://doi.org/10.1021/acs.est.1c08673

Pruden, A., Arabi, M. & Storteboom, H. N. (2012). Correlation between upstream human activities and riverine antibiotic resistance genes. *Environmental Science and Technology*, 46, 11541–11549. https://doi.org/10.1021/es302657r

Quintela-Baluja, M., Abouelnaga, M., Romalde, J., Su, J. Q., Yu, Y., Gomez-Lopez, M., Smets, B., Zhu, Y. G. & Graham, D. W. (2019). Spatial ecology of a wastewater network defines the antibiotic resistance genes in downstream receiving waters. *Water Research*, 162, 347–357. https://doi.org/10.1016/j.watres.2019.06.075

Quintela-Baluja, M., Frigon, D., Abouelnaga, M., Jobling, K., Romalde, J. L., Gomez Lopez, M. & Graham, D. W. (2021). Dynamics of integron structures across a wastewater network – Implications to resistance gene transfer. *Water Research*, 206. https://doi.org/10.1016/j.watres.2021.117720

Rice, E. W., Allen, M. J., Brenner, D. J. & Edberg, S. C. (1991). Assay for β-D-glucuronidase in species of the genus Escherichia and its applications for drinking-water analysis. *Applied and Environmental Microbiology, 57*, 592-593. https://doi.org/10.1128/aem.57.2.592-593.1991

Riley, L. W. (2014). Pandemic lineages of extraintestinal pathogenic *Escherichia coli*. *Clinical Microbiology and Infection*. 20, 380–390. https://doi.org/10.1111/1469-0691.12646

Robins, K., Leonard, A. F. C., Farkas, K., Graham, D. W., Jones, D. L., Kasprzyk-Hordern, B., Bunce, J. T., Grimsley, J. M. S., Wade, M. J., Zealand, A. M. & McIntyre-Nolan, S. (2022). Research needs for optimising wastewater-based epidemiology monitoring for public health protection. *Journal of Water and Health*, 20, 1284–1313. https://doi.org/10.2166/wh.2022.026

Roer, L., Hansen, F., Frølund Thomsen, M. C., Knudsen, J. D., Hansen, D. S., Wang, M., Samulioniené, J., Justesen, U. S., Røder, B. L., Schumacher, H., Østergaard, C., Andersen, L. P., Dzajic, E., Søndergaard, T. S., Stegger, M., Hammerum, A. M. & Hasman, H. (2017). WGS-based surveillance of third-generation cephalosporin-resistant *Escherichia coli* from bloodstream infections in Denmark. *Journal of Antimicrobial Chemotherapy*, *72*, 1922–1929. https://doi.org/10.1093/jac/dkx092

Rojas-Jiménez, J., Jiménez-Pearson, M. A., Duarte-Martínez, F., Brenes-Mora, E., Arguedas, R. & Barquero-Calvo, E. (2022). First Report of a Multidrug-Resistant ST58 Escherichia coli Harboring Extended-Spectrum Beta-Lactamase of the CTX-M-1 Class in a Fecal Sample of a Captive Baird's Tapir (*Tapirus bairdii*) in Costa Rica, Central America. *Microbial Drug Resistance*, 28, 143–148. https://doi.org/10.1089/mdr.2020.0339

Rompré, A., Servais, P., Baudart, J., de-Roubin, M. R. & Laurent, P. (2002). Detection and enumeration of coliforms in drinking water: current methods and emerging approaches. *Journal of Microbiological Methods*, 49, 31–54. https://doi.org/10.1016/s0167-7012(01)00351-7

Rowe, W. P. M., Baker-Austin, C., Verner-Jeffreys, D. W., Ryan, J. J., Micallef, C., Maskell, D. J. & Pearce, G. P. (2017). Overexpression of antibiotic resistance genes in hospital effluents over time. *Journal of Antimicrobial Chemotherapy*, 72, 1617–1623. https://doi.org/10.1093/jac/dkx017

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

Smith, K. A. (2012). Louis Pasteur, the father of immunology? *Frontiers in Immunology*, 3, Article 68. https://doi.org/10.3389/fimmu.2012.00068

Stoesser, N., Sheppard, A. E., Pankhurst, L., De Maio, N., Moore, C. E., Sebra, R., Turner, P., Anson, L. W., Kasarskis, A., Batty, E. M., Kos, V., Wilson, D. J., Phetsouvanh, R., Wyllie, D., Sokurenko, E., Manges, A. R., Johnson, T. J., Price, L. B., Peto, T. E., Johnson, J. R., Didelot, X., Walker, A. S. & Crook, D. W. (2016). Modernizing Medical Microbiology Informatics Group

(MMMIG). Evolutionary History of the Global Emergence of the *Escherichia coli* Epidemic Clone ST131. *mBio*. 7. :e02162. doi: 10.1128/mBio.02162-15.

Stokes, H. W. & Gillings, M. R. (2011). Gene flow, mobile genetic elements and the recruitment of antibiotic resistance genes into Gram-negative pathogens. *FEMS Microbiology Reviews* 35, 790–819. https://doi.org/10.1111/j.1574-6976.2011.00273.x

Stosic, M. S., Sunde, M., Mo, S. S., Telke, A. A. & Rudi, K. (2021). Interference of ISEcp1-blaCTX-M-1 with the shufflon rearrangement in IncI1 plasmids. *Plasmid*, 116. https://doi.org/10.1016/j.plasmid.2021.102578

Subirats, J., Triadó-Margarit, X., Mandaric, L., Acuña, V., Balcázar, J. L., Sabater, S. & Borrego, C. M. (2017). Wastewater pollution differently affects the antibiotic resistance gene pool and biofilm bacterial communities across streambed compartments. *Molecular Ecology*, 26, 5567–5581. https://doi.org/10.1111/mec.14288

Suzuki, S., Pruden, A., Virta, M. & Zhang, T. (2017). Editorial: Antibiotic resistance in aquatic systems. *Frontiers in Microbiology*. 8. https://doi.org/10.3389/fmicb.2017.00014

Suzuki, Y., Hashimoto, R., Xie, H., Nishimura, E., Nishiyama, M., Nukazawa, K. & Ishii, S. (2019). Growth and antibiotic resistance acquisition of *Escherichia coli* in a river that receives treated sewage effluent. *Science of the Total Environment*, 690, 696–704. https://doi.org/10.1016/j.scitotenv.2019.07.050

Szczepanowski, R., Braun, S., Riedel, V., Schneiker, S., Krahn, I., Pühler, A. & Schlüter, A. (2005). The 120 592 bp IncF plasmid pRSB107 isolated from a sewage-treatment plant encodes nine different antiobiotic-resistance determinants, two iron-acquisition systems and other putative virulence-associated functions. *Microbiology*, 151, 1095–1111. https://doi.org/10.1099/mic.0.27773-0

Tooke, C. L., Hinchliffe, P., Bragginton, E. C., Colenso, C. K., Hirvonen, V. H. A., Takebayashi, Y. & Spencer, J. (2019). β-Lactamases and β-Lactamase Inhibitors in the 21st Century. *Journal of Molecular Biology*, 431, 3472–3500. https://doi.org/10.1016/j.jmb.2019.04.002

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

Varani, A., He, S., Siguier, P., Ross, K. & Chandler, M. (2021). The IS6 family, a clinically important group of insertion sequences including IS26. *Mobile DNA*, 12. https://doi.org/10.1186/s13100-021-00239-x

Vila, J., Sáez-López, E., Johnson, J. R., Römling, U., Dobrindt, U., Cantón, R., Giske, C. G., Naas, T., Carattoli, A., Martínez-Medina, M., Bosch, J., Retamar, P., Rodríguez-Bañõ, J., Baquero, F. & Soto, S. M. (2016). *Escherichia coli*: An old friend with new tidings. *FEMS Microbiology Reviews* 40, 437–463. https://doi.org/10.1093/femsre/fuw005

Villa, L., García-Fernández, A., Fortini, D. & Carattoli, A. (2010). Replicon sequence typing of IncF plasmids carrying virulence and resistance determinants. *Journal of Antimicrobial Chemotherapy*, 65, 2518–2529. https://doi.org/10.1093/jac/dkq347

WHO - (2021) -Tripartite AMR country self-assessment survey (TrACSS) 2019-2020. (n.d.).

WHO integrated global surveillance on ESBL-producing *E. coli* using a "One Health" approach: Implementation and opportunities. (2021). http://apps.who.int/bookorders.

Woodford, N., Carattoli, A., Karisik, E., Underwood, A., Ellington, M. J. & Livermore, D. M. (2009). Complete nucleotide sequences of plasmids pEK204, pEK499, and pEK516, encoding CTX-M enzymes in three major *Escherichia coli* lineages from the United Kingdom, all belonging to the international O25:H4-ST131 clone. *Antimicrobial Agents and Chemotherapy*, 53, 4472–4482. https://doi.org/10.1128/AAC.00688-09

Woodford, N., Turton, J. F. & Livermore, D. M. (2011). Multiresistant Gram-negative bacteria: The role of high-risk clones in the dissemination of antibiotic resistance. *FEMS Microbiology Reviews*, 35, 736–755. https://doi.org/10.1111/j.1574-6976.2011.00268.

Wright, G. D. (2019). Environmental and clinical antibiotic resistomes, same only different. *Current Opinion in Microbiology,* 51, 57-63. https://doi.org/10.1016/j.mib.2019.06.005

Wyness, A. J., Paterson, D. M., Defew, E. C., Stutter, M. I. & Avery, L. M. (2018). The role of zeta potential in the adhesion of *E. coli* to suspended intertidal sediments. *Water Research*, 142, 159–166. https://doi.org/10.1016/j.watres.2018.05.054

Wyrsch, E. R., Chowdhury, P. R., Wallis, L., Cummins, M. L., Zingali, T., Brandis, K. J. & Djordjevic, S. P. (2020). Whole-genome sequence analysis of environmental *Escherichia coli* from the faeces of straw-necked ibis (*Threskiornis spinicollis*) nesting on inland wetlands. *Microbial Genomics*, 6, 1–16. https://doi.org/10.1099/mgen.0.000385

Xiong, R., Liu, Y., Pu, J., Liu, J., Zheng, D., Zeng, J., Chen, C., Lu, Y. & Huang, B. (2021). Indole Inhibits IncP-1 Conjugation System Mainly Through Promoting korA and korB Expression. *Frontiers in Microbiology*, 12. https://doi.org/10.3389/fmicb.2021.628133

Zarkan, A., Liu, J., Matuszewska, M., Gaimster, H. & Summers, D. K. (2020). Local and Universal Action: The Paradoxes of Indole Signalling in Bacteria. *Trends in Microbiology* 28, 566–577. https://doi.org/10.1016/j.tim.2020.02.007

Zhang, J., Chen, L., Wang, J., Butaye, P., Huang, K., Qiu, H., Zhang, X., Gong, W. & Wang, C. (2018). Molecular detection of colistin resistance genes (mcr-1 to mcr-5) in human vaginal swabs. *BMC Research Notes*, 11. https://doi.org/10.1186/s13104-018-3255-3

Zhang, Y., Gu, A. Z., He, M., Li, D. & Chen, J. (2017). Subinhibitory Concentrations of Disinfectants Promote the Horizontal Transfer of Multidrug Resistance Genes within and across Genera. *Environmental Science and Technology*, 51, 570–580. https://doi.org/10.1021/acs.est.6b03132

Zhao, R., Feng, J., Liu, J., Fu, W., Li, X. & Li, B. (2019). Deciphering of microbial community and antibiotic resistance genes in activated sludge reactors under high selective pressure of different antibiotics. *Water Research*, 151, 388–402. https://doi.org/10.1016/j.watres.2018.12.034

Zhao, X., Zhao, H., Zhou, Z., Miao, Y., Li, R., Yang, B., Cao, C., Xiao, S., Wang, X., Liu, H., Wang, J. & Yang, Z. (2022). Characterization of Extended-Spectrum β-Lactamase-Producing *Escherichia coli* Isolates That Cause Diarrhea in Sheep in Northwest China . *Microbiology Spectrum*, 10. https://doi.org/10.1128/spectrum.01595-22

Zhou, Z. C., Feng, W. Q., Han, Y., Zheng, J., Chen, T., Wei, Y. Y., Gillings, M., Zhu, Y. G. & Chen, H. (2018). Prevalence and transmission of antibiotic resistance and microbiota between humans and water environments. *Environment International*, 121, 1155–1161. https://doi.org/10.1016/j.envint.2018.10.032

Zhou, Z., Alikhan, N.-F., Mohamed, K., Fan, Y., Study Group, A., Achtman, M., Brown, D., Petrovska, L., Davies, R., Robertson, A., Weill, X. & Accou-Demartin, M. (n.d.). The EnteroBase User's Guide, with case studies on Salmonella transmissions, *Yersinia pestis* phylogeny and *Escherichia* core genomic diversity. Retrieved from https://github.com/zheminzhou/EToKi

Zingali, T., Reid, C. J., Chapman, T. A., Gaio, D., Liu, M., Darling, A. E. & Djordjevic, S. P. (2020). Whole genome sequencing analysis of porcine faecal commensal *Escherichia coli* carrying class 1 integrons from sows and their offspring. *Microorganisms*, 8. https://doi.org/10.3390/microorganisms8060843

Zong, Z., Ginn, A. N., Dobiasova, H., Iredell, J. R. & Partridge, S. R. (2015). Different IncI1 plasmids from *Escherichia coli* carry ISEcp1-blaCTX-M-15 associated with different Tn2-derived elements. *Plasmid*, 80, 118–126. https://doi.org/10.1016/j.plasmid.2015.04.007