

Escherichia coli in urban marine sediments: interpreting virulence, biofilm formation, halotolerance, and antibiotic resistance to infer contamination or naturalization

Isabel K. Erb^{1,2}, Carolina Suarez³, Ellinor M. Frank^{2,3}, Johan Bengtsson-Palme^{4,5,6}, Elisabet Lindberg⁷, Catherine J. Paul^{1,3,*}

¹Applied Microbiology, Department of Chemistry, Lund University, PO Box 124, SE-221 00 Lund, Sweden

²Sweden Water Research AB, Ideon Science Park, Scheelevägen 15, SE-223 70 Lund, Sweden

³Water Resources Engineering, Department of Building and Environmental Technology, Lund University, PO Box 118, SE-221 00 Lund, Sweden

⁴Division for Systems and Synthetic Biology, Department of Life Sciences, SciLifeLab, Chalmers University of Technology, Kemivägen 10, SE-412 96 Gothenburg, Sweden

⁵Department of Infectious Diseases, Institute of Biomedicine, The Sahlgrenska Academy, University of Gothenburg, Guldhedsgatan 10, SE-413 46 Gothenburg, Sweden

⁶Centre for Antibiotic Resistance research (CARE), SE-413 45 Gothenburg, Sweden

⁷City of Helsingborg, Department of City Planning, Järnvägsgatan 22, SE-252 25 Helsingborg, Sweden

*Corresponding author. Applied Microbiology, Department of Chemistry, Lund University, PO Box 124, SE-221 00 Lund, Sweden. E-mail: catherine.paul@tvrl.lth.se
Editor: [Kimberly Kline]

Abstract

Marine sediments have been suggested as a reservoir for pathogenic bacteria, including *Escherichia coli*. The origins, and properties promoting survival of *E. coli* in marine sediments (including osmotolerance, biofilm formation capacity, and antibiotic resistance), have not been well-characterized. Phenotypes and genotypes of 37 *E. coli* isolates from coastal marine sediments were characterized. The isolates were diverse: 30 sequence types were identified that have been previously documented in humans, livestock, and other animals. Virulence genes were found in all isolates, with more virulence genes found in isolates sampled from sediment closer to the effluent discharge point of a wastewater treatment plant. Antibiotic resistance was demonstrated phenotypically for one isolate, which also carried tetracycline resistance genes on a plasmid. Biofilm formation capacity varied for the different isolates, with most biofilm formed by phylogroup B1 isolates. All isolates were halotolerant, growing at 3.5% NaCl. This suggests that the properties of some isolates may facilitate survival in marine environments and can explain in part how marine sediments can be a reservoir for pathogenic *E. coli*. As disturbance of sediment could resuspend bacteria, this should be considered as a potential contributor to compromised bathing water quality at nearby beaches.

Keywords: environmental *Escherichia coli*; biofilm; halotolerance; whole genome sequencing; antibiotic resistance; bathing water

Abbreviations

ANI:	Average nucleotide identity
ARG:	Antibiotic resistance gene
DP:	Discharge point
<i>E. coli</i> :	<i>Escherichia coli</i>
ExPEC:	Extraintestinal pathogenic <i>Escherichia coli</i>
GPS:	Global positioning system
LB:	Lysogeny broth
PCR:	Polymerase chain reaction
ST:	Sequence type
WGS:	Whole genome sequence
W-UPEC:	Wastewater-urinary pathogenic <i>Escherichia coli</i>
WWTP:	Wastewater treatment plant

Introduction

Escherichia coli is a facultative-anaerobic Gram-negative bacteria of the vertebrate gut, and detection of this bacteria is frequently

used as indicator for fecal contamination (Odonkor and Ampofo 2013), and thus the presence of viruses, bacteria, or parasites which can cause serious diseases (Whitman et al. 2014). In designated swimming areas like beaches, monitoring programs for *E. coli* and other indicators are essential to minimize the health risks to the public (Halliday and Gast 2011). While *E. coli* itself can cause a variety of different diseases such as enteritis, urinary tract infections, and wound infections (Kaper et al. 2004), *E. coli* is also a commensal bacteria of warm-blooded animals and its presence is not necessarily connected to a direct risk to human health (Tenaillon et al. 2010). Understanding the origins of *E. coli* in the environment is thus an important factor in assessing both actual health risks and identify possible sources of contamination.

Contamination of urban bathing water with *E. coli* is often associated with multiple sources, such as anthropogenic activities, stormwater runoff, wastewater discharge, and animals in the marine environment (Haile et al. 1999, Solo-Gabriele et al. 2000, Alves et al. 2014). *Escherichia coli* has been detected in sediments and in

Received 23 January 2024; revised 28 June 2024; accepted 13 August 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of FEMS. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

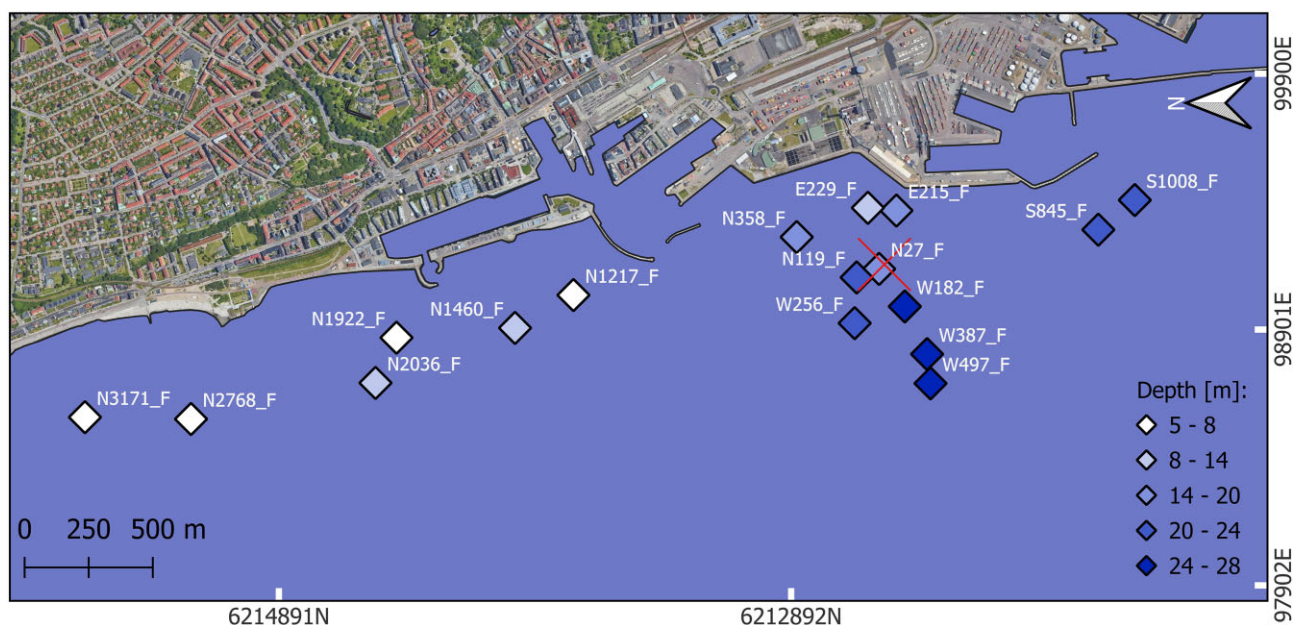


Figure 1. Sampling locations. The filled colour in each square corresponds to the sampling depth in meters. The approximate location of the WWTP effluent DP is at location N27_F and marked with a cross. The initial letter in the sample name indicates its position relative to the DP of the WWTP with N for North, E for east, S for south, and W for west, while the following number indicates the distance to the DP in meters.

beach sand (Halliday and Gast 2011, Pachepsky and Shelton 2011, Vignaroli et al. 2013, Salam et al. 2021). These sediments may be a source of *E. coli* (Ishii et al. 2007) and in freshwater environments, resuspension from sediments has been shown to contribute *E. coli* into the water phase (Schang et al. 2018). A previous study suggested that marine sediments in the proximity of an effluent discharge point (DP) of a wastewater treatment plant (WWTP) could serve as a potential reservoir for *E. coli* (Frank et al. 2024).

There is growing evidence that *E. coli* can persist and sometimes even thrive in environments other than in warm-blooded animals, where they are normally found. For example, *E. coli* belonging to phylogroup B1 are commonly observed in water environments (Touchon et al. 2020). In these environments, organic matter might be a sufficient nutrient source and facilitates survival of *E. coli* (Gerba and McLeod 1976). While recent studies have characterized both genotypes and phenotypes of environmental *E. coli* isolates (Julian et al. 2015, Kindle et al. 2019), including marine environments (Vignaroli et al. 2013, Grunwald et al. 2022), there is limited knowledge about *E. coli* from marine sediments. Additionally, it is unclear whether *E. coli* are still suitable indicators of fecal pollution, and if they pose potential risk to human health.

One possibility is that urban marine sediments select for *E. coli* strains that possess properties that facilitate survival and thus lead to naturalization. This should also mean that occurrence of virulence, antibiotic resistance, genetic determinants, and phylogroup of whole genome sequences (WGSs) will not be random and depend instead on the proximity of the isolates' location to the WWTP DP.

This study used whole genome sequencing and standard phenotypic assays to address this possibility. Thirty-seven *E. coli* isolates were obtained from marine sediments. Genes associated with virulence, biofilm formation capacity, and antibiotic resistance were compared to phenotypic determination of antibiotic resistance profiles, biofilm formation capacity, and halotolerance. These phenotypes were selected to examine if the isolates were able to survive in sediments due to adaptation to the specific en-

vironment, and if these characteristics could be explained by genetic determinants. In addition, by examining both genotype and phenotype, the suitability of DNA-based methods to determine health risks, including antibiotic resistance could be assessed. Lastly, to examine possible sources for the *E. coli* present in the sediments, whole genome sequences were compared to sequences of *E. coli* with defined origins.

Material and methods

Sampling

Sediments were retrieved using a core sampler (diameter: 15 cm) from 17 locations in the Öresund near the city of Helsingborg, Sweden on 24th August 2021. All samples not located in proximity to the DP of the WWTP were sampled perpendicular to the bathing locations on the coast ~300 m from the land (Table S4, Fig. 1). The top 1 cm of sediment from each core was collected into sterile 50 ml centrifuge tubes. The samples were kept on ice during transport back to the laboratory (maximum time: 4 h). Location names, global positioning system (GPS) coordinates, and sampling depth can be found in Table S4 in the supporting materials. The map of the sampling locations was generated in QGIS 3.22.6 (QGIS Development Team 2023) using the coordinate reference system SWEREF99 13.30 (EPSG:5847).

Quantification and isolation of *E. coli*

In the laboratory, sediments were transferred into sterile Petri dishes and mixed, then 10 ml of each sample was transferred into a new centrifuge tube. Milli Q-water was added to a total volume of 30 ml and filled tubes were gently mixed for 18 h at 5–6°C on a rocking platform. The samples were then allowed to settle for 1 h before the top 11 ml, maximum amount of water without any dark sediment particles coming along, was poured off into a measuring cylinder. Water from duplicates was pooled (total volume of 22 ml) and diluted with Milli Q-water to generate 5x and 50x dilutions of

the water extracted from the sediments. *Escherichia coli* concentrations were determined using Colilert-18 and Quanti-Tray/2000 (IDEXX, ME, USA). Concentrations were multiplied by 5 to obtain MPN/100 ml sediment.

To isolate individual *E. coli*, positive wells in the Quanti-Tray/2000 wells from either 5x or 50x duplicates were randomly punctured with a sterile needle. Approximately, 200 μ l liquid was transferred to HICrome Agar B plates (Sigma-Aldrich, MA, USA) and streaked for single colonies. Plates were incubated for 14 h at 36°C and individual blue colonies, indicating presumptive *E. coli*, were streaked again before being grown in liquid lysogeny broth (LB). For each of the 15 locations where *E. coli* was observed, at least two isolates were selected and archived in 25% glycerol at -80°C. Overnight cultures were grown in liquid LB from single colonies streaked on LB agar plates for additional experiments and DNA isolation.

DNA extraction and sequencing

DNA was extracted from overnight cultures using the GeneJet Genomic DNA Purification Kit (Thermo Fisher Scientific, MA, USA) according to the manufacturer's instructions for Gram-negative bacteria. DNA concentrations were determined by Qubit dsDNA HS Assay kit (Thermo Fisher Scientific). Forty-one isolates were selected for sequencing. Libraries were prepared with Illumina DNA prep (M) tagmentation. A NextSeq 550 (Illumina, CA, USA) was used for 2 \times 150 bp paired-end sequencing with a NextSeq 500/550 Mid Output Kit v2.5 (Illumina).

Genome analysis

Reads were assembled with SKESA v2.2 (Souvorov et al. 2018), cross-species and intraspecies contamination assessed with CheckM v1.0.12 (Parks et al. 2015) and ConFindr v0.7.1 (Low et al. 2019). This led to exclusion of two genomes with sequence contamination. Taxonomic affiliation was estimated with GTDB-Tk v2.1.0 (Chaumeil et al. 2019) using the GTDB Release 07-RS207 taxonomy (Parks et al. 2022). An additional two genomes, classified as *Shewanella algae* and *Serratia ureilytica*, were excluded from the dataset. The remaining 37 genomes were classified as *E. coli*.

Phylogenetic groups (phylogroups) were estimated with ClermonTyping (Beghain et al. 2018). *Escherichia coli* sequence types (STs) were determined with the multilocus sequence typing (MLST) MLST 2.0 webtool (Larsen et al. 2012) using the Achtman scheme (Wirth et al. 2006). Average nucleotide identity (ANI) of the *E. coli* genomes was estimated with fastANI v1.33 (Jain et al. 2018). ANI similarity values were converted to dissimilarities, and then a dendrogram was made in R by hierarchical clustering with average linkage method.

Genes potentially linked to antibiotic resistance in the assembled genomes were identified with the Resistance Gene Identifier v6.0.2 using the Comprehensive Antibiotic Resistance Database v3.2.6 as reference (Alcock et al. 2019, 2023). Default settings for detection of only perfect and strict hits were used. Genes associated with virulence in *E. coli* were identified with VirulenceFinder 2.0 (Joensen et al. 2014, Malberg Tetzschner et al. 2020) with the 2022-02-12 database version, using a BLAST search of the assembled genomes with a 90% threshold identity.

The web tool MLST query of Enterobase v1.1.3 (Zhou et al. 2020) with the Achtman scheme was used to examine where STs had been previously observed. Different individual sources were pooled into main source groups (Table S1). A heatmap showing

observed prevalence of sources of the different STs was created using R version 4.1.1 and the R package heatmaply.

Genes linked to biofilm production and halotolerance were annotated in assembled genomes with eggNOG-mapper v2.1.9 (Cantalapiedra et al. 2021) using DIAMOND (Buchfink et al. 2021), the eggNOG 5.0 database (Huerta-Cepas et al. 2019), and Prodigal (Hyatt et al. 2010). Genes associated with biofilm production (Table S2) and osmotic stress regulation (Table S3) were identified by literature search for use in the presence and absence analysis.

Disk diffusion assays were performed according to the antimicrobial susceptibility testing disk diffusion method version 10.0, January 2022, of the European committee of antimicrobial susceptibility testing (EUCAST). Isolates were grown on Müller-Hinton agar plates and tested for their susceptibility with antibiotic disks (Oxoid™ antimicrobial susceptibility disks, Thermo Fisher Scientific) on cefotaxime (5 μ g), meropenem (10 μ g), ciprofloxacin (5 μ g), tetracycline (30 μ g), and gentamicin (10 μ g).

Biofilm test

Biofilm formation capacity was assessed as previously described (Coffey and Anderson 2014) with a few modifications. LB and LB with 3.5% NaCl were used as growth medium. 3.5% NaCl was chosen since this concentration approximately resembles the maximum NaCl concentration found in the marine sediments, where isolates were sampled from (Leppranta and Myrberg 2009).

The 96-well plates were incubated at 37°C with shaking at 900 rpm every 5 min in a microplate reader (Multiskan Ascent, Thermo Fisher Scientific). The absorbance was measured for triplicates at 620 nm every 30 min for 8 h before staining with crystal violet. Crystal violet intensity was quantified using the microplate reader, measuring the absorbance at 550 nm. LB media was used as a negative control.

Salinity test

The same experimental setup was used as described for the biofilm test, without crystal violet staining.

Data describing the growth in LB and LB with 3.5% NaCl was analysed in R v4.1.1 (R Core Team 2021) using the R package Growthcurver v0.3.1. with the background correction 'min' (Sprouffske and Wagner 2016).

Statistical tests

All statistical comparisons were performed in R version 4.1.1 (R Core Team 2021). Differences in biofilm formation capacity and generation time in LB and LB with 3.5% NaCl were identified using the Wilcoxon signed-ranked test. One-way ANOVA and Tukey HSD tests were performed to show differences in biofilm formation capacity, abundance of virulence genes, and generation time between the phylogroups.

To identify trends in the distribution of genes among phylogroups, and to explore potential links with phenotype, redundancy analyses (RDA) were done in the R package Vegan v 2.6.4. A matrix of the number of observed genes across the *E. coli* genomes was made from: the (1) genes associated with biofilm formation according to the literature survey, or (2) all the KEGG orthologs from the eggNOG annotation, except for core genes. These were used as input to the RDA. The constraining variables were generation time in LB with and without NaCl, and absorbance of the biofilm assay with and without NaCl. A Hellinger transformation was used for the gene matrices, and the constraining variables were standardized.

Results and discussion

Viable *E. coli* was observed in coastal marine sediments

The presence of viable *E. coli* in coastal marine sediments was evaluated for 17 locations alongside the coastline of the city of Helsingborg, Sweden (Fig. 1). The concentration of *E. coli* in surface sediments ranged from 0 to 4540 MPN/100 ml, with an average of 1113 MPN/100 ml (Table S4). *Escherichia coli* was observed in 15 out of 17 sampling locations. For each of the 15 locations where *E. coli* was observed, at least two isolates were selected for further analysis.

While *E. coli* is used as an indicator of fecal contamination, under the assumption that *E. coli* is associated with the gut microbiota and short-lived in the environment, recent studies have reported survival and growth in the environment (Luo et al. 2011, van Elsas et al. 2011, Rumball et al. 2021). Environmental *E. coli* have been isolated from diverse environments like groundwater (Tropea et al. 2021), seawater, shellfish (Balière et al. 2015), and freshwater sand (Walk et al. 2007). This suggests that while some of the observed *E. coli* in this study could constitute recent contamination, adaptation to natural environments (i.e. naturalization) of some populations was also considered. WWTPs play a crucial role in contaminating water environments with *E. coli* (Anastasi et al. 2012, Zieliński et al. 2021), and the WWTP in this study has been proposed as source of contamination near the DP (Frank et al. 2024). Therefore, further characterization of isolates using whole genome sequencing was pursued in an attempt to further clarify the source of the bacteria.

Whole genomes reveal diversity of *E. coli* in coastal marine sediments

Whole genome sequencing was performed for 41 isolates, followed by genome assembly. Thirty-seven isolates were classified as *E. coli*, with genome completeness of $\geq 99.32\%$ and minimal contamination of foreign DNA ($\leq 1.56\%$). Genome size of these 37 genomes ranged between 4.50 and 5.11 Mbp, with an average of 4.84 Mbp (Table S5). *In silico* multiplex polymerase chain reaction (PCR) (Beghain et al. 2018), assigned 36 of the 37 *E. coli* isolates to phylogroups A, B1, B2, D, and E and for 35 of the 37 isolates there was a match between *in silico* PCR and mash clustering results of ClermonTyping (Table S6). W387_F1 and E215_F2 isolates were assigned to phylogroup C and 'unknown', respectively, using *in silico* PCR, while mash assigned both isolates to phylogroup A. Similar results to the mash classification were observed when genomes were grouped using ANI estimations (Figure S1). MLST identified 30 different STs further demonstrating the high genetic diversity of *E. coli* in these coastal sediments (Table S6).

One isolate was assigned to phylogroup E, which includes the highly pathogenic serotype O157:H7, and many commensals (Clermont et al. 2021). Fifteen of 37 isolates were assigned to phylogroup B1. This phylogroup has been previously associated with animals (Higgins et al. 2007, Carlos et al. 2010, Johnson et al. 2017) but has also been frequently observed in water environments (Berthe et al. 2013, Touchon et al. 2020, Rumball et al. 2021). This suggests that some isolates in Öresund sediments may originate from animals or could occur naturally in the marine environment. Six isolates from sediments near the WWTP DP were identified as phylogroup B2 (Fig. 1, Figure S1). This phylogroup harbours several extraintestinal pathogenic *E. coli* (ExPEC) strains (Johnson and Russo 2002, Denamur et al. 2021). It has been shown that ExPEC can survive wastewater treatment (Raboni et al. 2016, Zhi et al. 2020, Yu et al. 2022). ExPEC often belong to the phylogroups B2

and D (Picard et al. 1999, Johnson and Russo 2002, Denamur et al. 2021). More specifically, Wastewater-urinary pathogenic *E. coli* (W-UPEC) cluster within the phylogroup B1, B2, and D, with B2 as the predominant group within W-UPEC (Zhi et al. 2020). Taken together, this indicates that some isolates, particularly those isolated from sediments in proximity to the WWTP could be potentially pathogenic and likely also originated from the WWTP, while others are more likely to have entered the environment from other sources including combined sewer overflows, storm water, and animals habituating the area (Wright et al. 2009, McCarthy et al. 2017, McGinnis et al. 2022). This was also supported by the diversity observed in the MLST analysis: a recent study in the marine environment in the Salish Sea demonstrated that high diversity correlated with areas that were partially impacted by wastewater outflows (Grunwald et al. 2022).

Information about source based on STs

EnteroBase can associate STs with different sources (Zhou et al. 2020) and was applied to identify the most likely sources associated with the STs assigned to the isolates (Fig. 2). STs in this study corresponded to STs that have been mainly documented in humans and animals, including livestock. Approximately 20% of the observed STs corresponded to STs frequently reported in humans, such as ST10, ST73, ST131, and ST127. ST131 and ST73 are among the major ExPEC strains and are considered pathogenic (Nicolas-Chanoine et al. 2014, Riley 2014). In addition, ST131 and ST127, both represented among the isolates, have been associated with wastewater (Finn et al. 2020, Zhi et al. 2020). ST131 has been observed in marine sediments (Vignaroli et al. 2013) and is associated with carriage of virulence factors causing complicated urinary tract infections with potential treatment failure (Can et al. 2015). ST8972, assigned to S1008_F1, has only been associated with two environmental isolates from surface soil (Dusek et al. 2018). ST214, which was assigned to N1922_F2, has only been observed in one *E. coli* isolated from humans (GenBank accession: GCA_900490405.1). Observations from a recent study showed that STs 10, 162, 362, and 2144 have been isolated from aquatic animals such as fish and seals (Grunwald et al. 2022).

Viable *E. coli* from sediments harbour a variety of virulence genes

Virulence genes were identified using VirulenceFinder 2.0. Sixty-six genes potentially linked to virulence were found in the genomes ($>90\%$ identity), with an average of 18.46 genes per isolate. Virulence gene patterns differed between the phylogroups, with B2 showing the highest abundance of virulence genes (Figure S3), consistent with what has been previously reported for phylogroup B2 (Picard et al. 1999). Among the observed genes of concern were genes encoding vacuolating autotransporter toxin (*vat*), fimbrial-like protein (*yfcV*), siderophore yersiniabactin receptor (*fyuA*), outer membrane hemin receptor (*chuA*), iron transport protein (*sitA*), increased serum survival lipoprotein (*Iss*), and outer membrane usher P fimbriae (*papC*), which have all been linked to virulence in ExPEC (Sarowska et al. 2019). These genes are also considered as predictor genes for uropathogenic potential (Spurbeck et al. 2012). This suggests that the phylogroup B2 isolates in this study are potentially ExPEC and uropathogenic. The prevalence of virulence genes can increase after wastewater treatment (Osińska et al. 2020), and the hypothesis that some of the *E. coli* isolates in this study originate from the WWTP is supported by the observation that the isolates from sediments in proximity

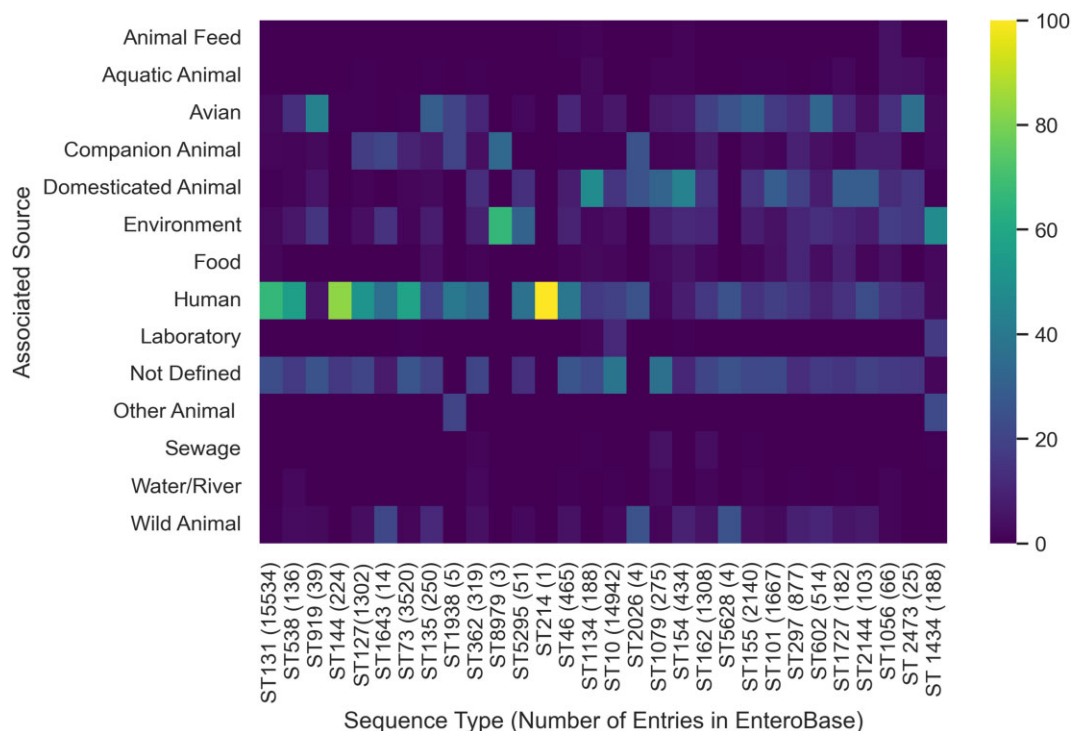


Figure 2. Percentage of entries in EnteroBase associated with different sources and the assigned STs of different isolates as of 19 December 2022. Colours in the heatmap correspond to the colour bar on the right. The corresponding isolates to the STs are displayed on the x-axis. Number of entries in EnteroBase for each ST is displayed in brackets next to STs.

of the DP carried more virulence genes than those at a greater distance (Fig. 3).

One of the genes coding for long polar fimbriae, *lpfA*, was only observed in 14 of the isolates identified as phylogroup B1 (Figure S2). The *lpfA* gene is common in B1 isolates, and encodes one of the major subunits of long polar fimbria (Madoshi et al. 2016, Zhou et al. 2021). This gene is often observed in intestinal pathogenic strains, including *E. coli* O157:H7 (Zhou et al. 2019) and has been associated with virulence via enhancement of adhesion and biofilm formation capacity (Ross et al. 2015). The ability to produce this fimbria might facilitate adhesion and biofilm formation capacity of *E. coli* in marine sediments, providing a survival advantage, and partially explaining why B1 phylogroup isolates were abundant.

Escherichia coli in coastal sediments are potential reservoirs for antibiotic resistance genes

Resistance to aminoglycosides, carbapenems, cephalosporins, fluoroquinolones, and tetracyclines was evaluated using the clinically accepted EUCAST protocol with the antibiotics cefotaxime (cephalosporin), meropenem (carbapenem), ciprofloxacin (fluoroquinolone), tetracycline, and gentamicin (aminoglycoside). The presence of genes in the isolate genomes associated with antibiotic resistance phenotypes was assessed using antibiotic resistance gene (ARG) prediction *in silico*. ARGs were observed in all genomes (Figures S4–S8).

All isolates were sensitive to all five antibiotics tested, with the exception of isolate N119_F1. This isolate was resistant to both ciprofloxacin and tetracycline (Tables S7–S11). The genome of isolate N119_F1 was the only isolate harbouring the genes *tet(B)* and *tetR* (strict hits) linked to tetracycline resistance (Figures S5 and S6) (Nguyen et al. 2014), and specific mutations within *gyrA* and

parC linked to fluoroquinolone resistance (strict hits) (Redgrave et al. 2014).

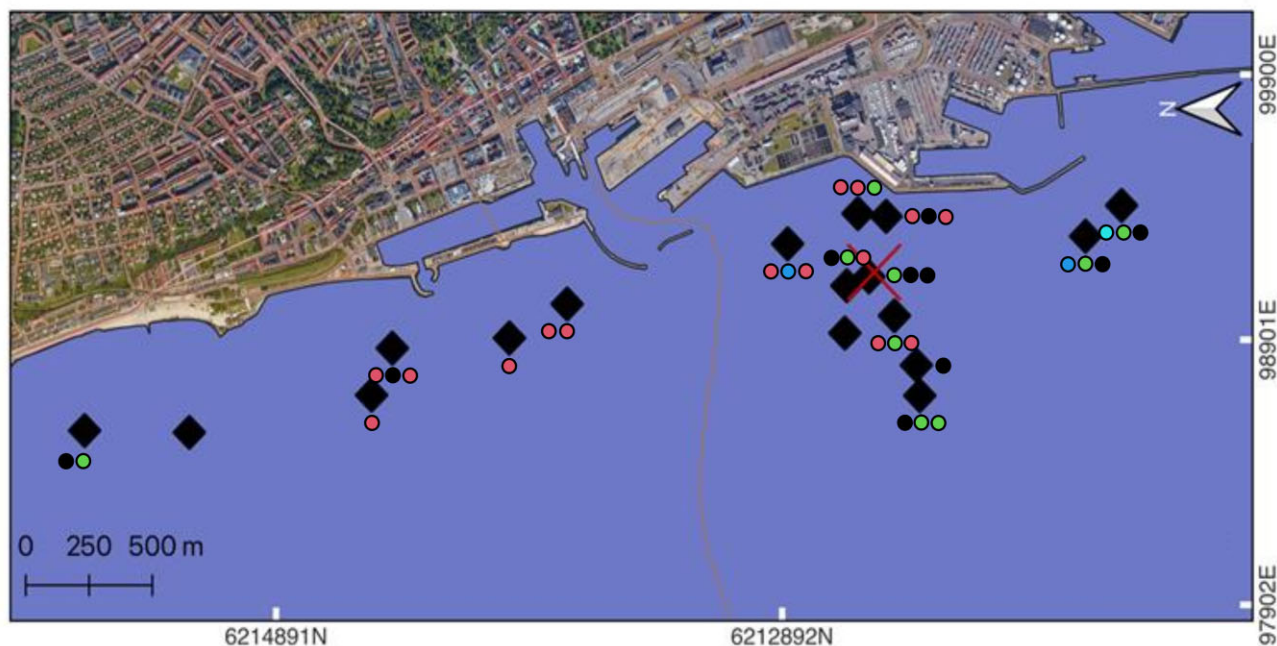
Further investigation of the genome obtained for N119_F1 identified a single contig predicted to be an IncR plasmid (100% identity). This contig harboured the *tet(B)* and *tet(R)* genes as well as the *aph(3'')-Ib* (*strA*) and *aph(6)-Id* (*strB*) genes encoding phosphotransferases linked to aminoglycoside resistance (Shaw et al. 1993). As an alternative approach to identify putative plasmids, *de novo* assembly of plasmids was performed for the DNA reads from all isolates. This approach recovered a putative plasmid from N119_F1 carrying the *tet(B)*, *tetR*, *aph(3'')-Ib*, and *aph(6)-Id* genes. This strongly suggests, that in N119_F1, these ARGs located on a plasmid were responsible for the observed tetracycline resistance phenotype. Using the same approach, a putative IncI1-I(Alpha) plasmid carrying both *aph(3'')-Ib* and *aph(6)-Id* was identified for the isolate N119_F3.

While only one isolate demonstrated antibiotic resistance as defined in the EUCAST protocol, the other isolates could still be sources for spread ARGs via horizontal gene transfer within the sediments (Wang et al. 2021) with the N119_F1 and N119_F3 isolates posing an even greater risk for spread, as they carried ARGs likely to be present on a plasmid. Sediments in aquatic environments have previously been identified as potential reservoirs for ARGs (Luo et al. 2010, Wu et al. 2021) and sediments in this marine environment of the Öresund also could serve as a reservoir.

Escherichia coli show potential adaptation to the marine sediment environment

As the ability to form biofilm could facilitate survival of *E. coli* in marine sediments, the isolates were assessed for their biofilm formation capacity in LB and LB with elevated NaCl of 3.5% using the crystal violet biofilm growth assay in 96-well microtiter plates.

(A)



(B)

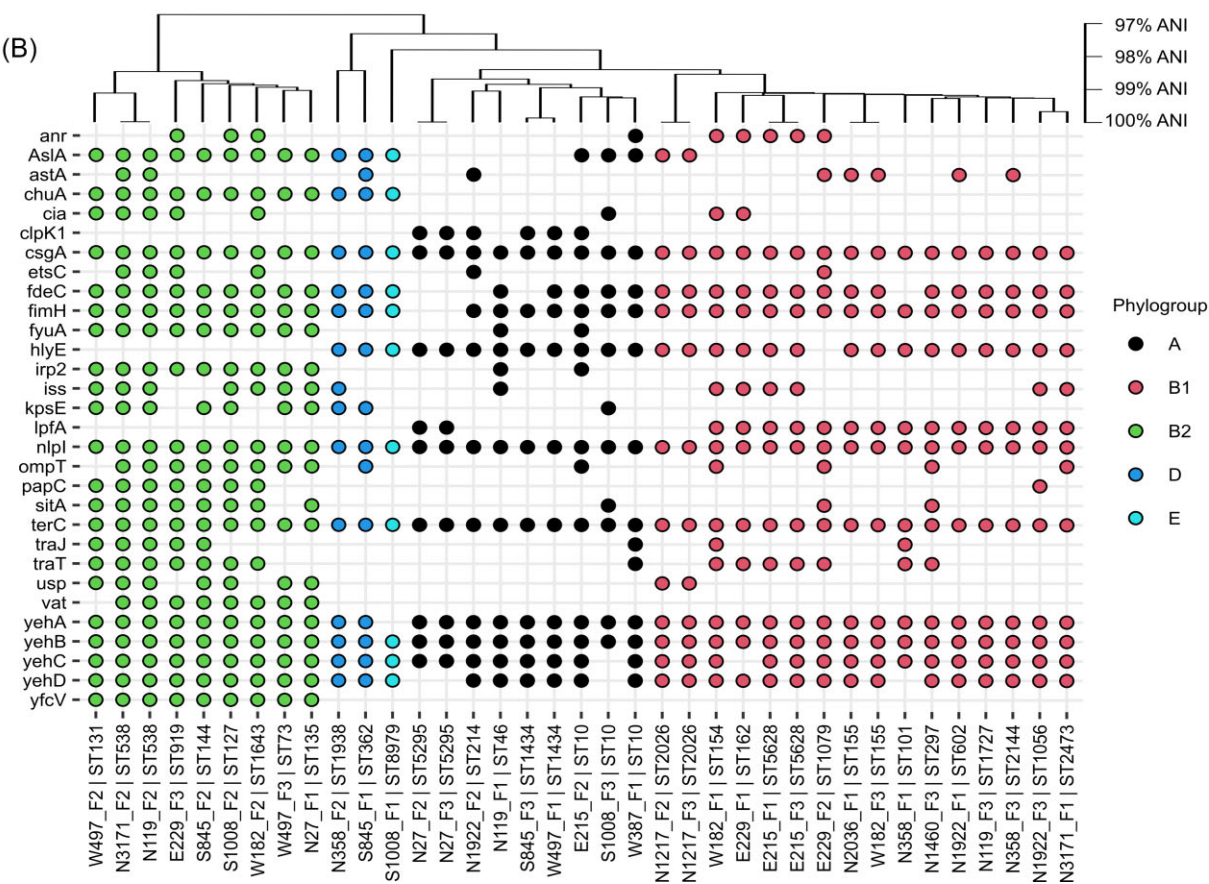


Figure 3. Geographical distribution, and virulence genes identified, in the different isolate phylogroups. (A) Distribution of assigned phylogroups of isolates in coastal sediment samples. DP of WWTP is marked with a cross. (B) Identification of genes potentially linked to virulence in the different isolates, with coloured dots indicating the gene and phylotype associated with the isolate named along the x-axis. To group isolates by their genome relatedness, isolates are clustered by ANI. Only genes detected in at least 20 isolates are shown. Less frequently detected virulence genes are described in Figure S2.

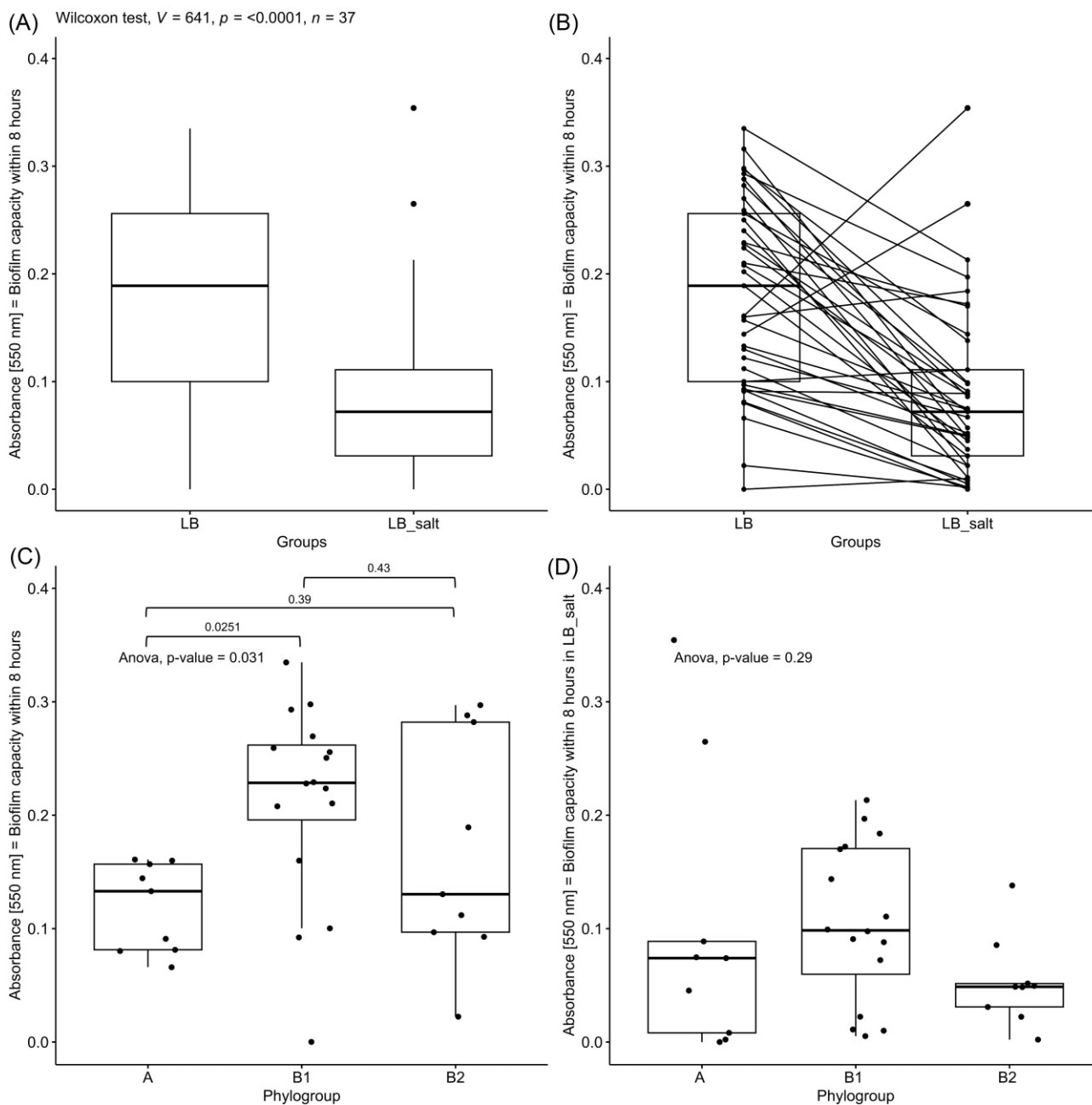


Figure 4. (A) Boxplot showing biofilm formation capacity of different isolates grown in LB and LB with 3.5% NaCl. Values above indicate the P -value of the Wilcoxon signed-rank exact test. (B) Linked boxplot of biofilm formation capacity of different isolates grown in LB and LB with 3.5% NaCl. (C) Boxplots of biofilm formation capacity in isolates grown in LB, grouped by assigned phylogroup A, B1, or B2 with ANOVA P -value (P -value = .031) and p -adj of the Tukey HDS test (above the square brackets). (D) Boxplots of biofilm formation capacity in isolates grown in LB with 3.5% NaCl, grouped by assigned phylogroup A, B1, or B2 with ANOVA P -value (P -value = .29).

For most isolates, biofilm formation at elevated NaCl was significantly lower. (Wilcoxon signed-rank exact test P -value = 1.79×10^{-6} ; Fig. 4A). Remarkably, isolate W497_F1 (ST1443) and S1008_F3 (ST10) showed increased biofilm formation capacity in LB with 3.5% salt (Fig. 4B). When grown in solely LB medium, abundant biofilm formation capacity ($OD_{550} > 0.3$) was observed for two isolates and 16 isolates had moderate biofilm formation capacity (OD_{550} 0.2–0.3). Minimal biofilm formation capacity was noticed for 10 isolates (OD_{550} 0.1–0.2) while nine isolates did not show growth as biofilm ($OD_{550} > 0.1$) (Figure S9A). Biofilm formation capacity was inhibited in most strains when grown in LB with 3.5% NaCl. In this condition, 27 strains did not show growth

as biofilm and 7 isolates showed minimal biofilm formation capacity. Moderate biofilm formation capacity was observed for two strains while abundant biofilm formation capacity was only observed in one strain (Figure S9B). Biofilm growth was associated with phylogroup when grown in LB (one-way ANOVA, P -value = .035; Fig. 4C), with phylogroup A forming less biofilm than phylogroup B1 (Tukey HDS, P -value = .022). No significant differences between phylogroups A and B2 (Tukey HDS, P -value = .34), and B1 and B2 (Tukey HDS, P -value = .47) were observed (Fig. 4). No significant differences in biofilm formation capacity between the phylogroups were observed when grown in LB medium with 3.5% NaCl (one-way ANOVA, P -value = .29).

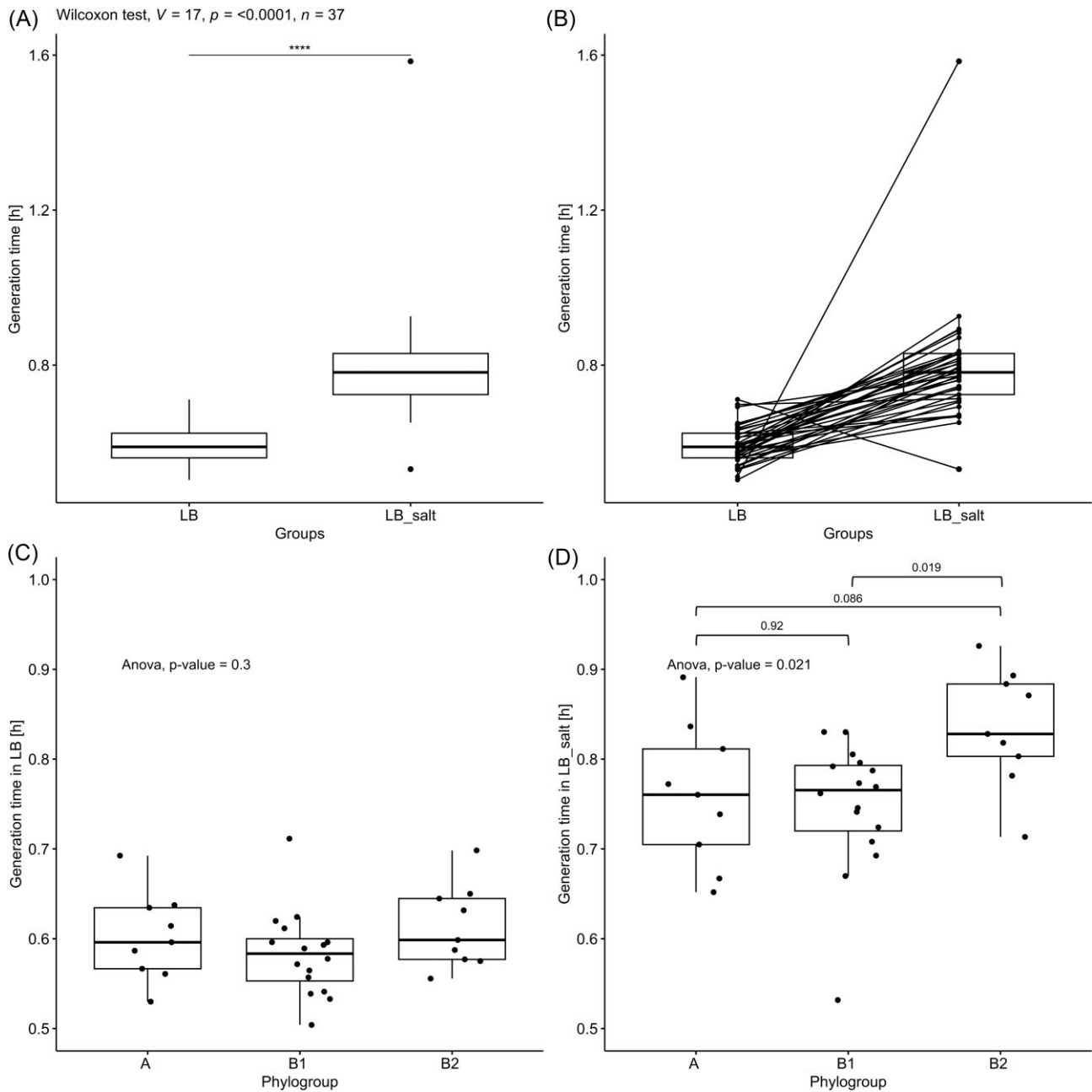


Figure 5. (A) Boxplot showing generation times of different isolates grown in LB and LB with 3.5% NaCl. Values above indicate the P-value of the Wilcoxon signed-rank exact test. (B) Linked boxplot of generation times of different isolates grown in LB and LB with 3.5% NaCl. (C) Boxplots of generation time in isolates grown in LB, grouped by assigned phylogroup A, B1, or B2 with ANOVA P-value (P-value = .3). (D) Boxplots of generation time in isolates grown in LB with 3.5% NaCl, grouped by assigned phylogroup A, B1, or B2 with ANOVA P-value (P-value = .021) and *p*-adj of the Tukey HSD test (above the square brackets).

Escherichia isolates assigned to phylogroup B1 tend to produce more biofilm than other phylogroups (Olowe et al. 2019), also observed in this study when comparing the B1 isolates to isolates from other phylogroups. As biofilm protects bacteria during growth on surfaces, this may facilitate colonization of sediments (Dang and Lovell 2016, Flemming and Wuertz 2019) and could explain why phylogroup B1 is often observed in water environments. Interestingly, for B1 isolate N2036_F1 (ST155), production of biofilm could not be detected, while for the closely related isolate W182_F3 (ST155) (ANI > 99.9%) biofilm growth was low in LB (Figure S9A). Isolates from phylogroup A had an overall lower level of biofilm formation capacity, which has previously been reported for this phylogroup (Martínez et al. 2006). Isolates from phylogroup A often lack several genes associated with biofilm pro-

duction (Figure S9C). Despite the overall low biofilm formation capacity in phylogroup A, S1008_F3 (ST10) and W497_F1 (ST1443) showed increased biofilm formation capacity when grown in LB with 3.5% NaCl compared when grown in LB alone.

Some trends between individual phenotypes of biofilm formation capacity and genome content could be identified (Figure S10). Isolate E229_F2 (ST1079) had the highest degree of biofilm formation capacity and possessed all analysed genes associated with biofilm grown in LB. With elevated NaCl, this isolate had lower but still moderate biofilm formation capacity. It had six copies of the *flu* (Antigen 43) gene, which has been linked to enhanced biofilm production in *E. coli* (van der Woude and Henderson 2008). S1008_F3 (ST10) and W497_F1 (ST1434) had all analysed genes, although biofilm formation capacity in LB was low, when grown with

3.5% NaCl moderate and high biofilm formation capacity, respectively was observed. N2036_F1 (ST155) and W182_F3 (ST 155) had no biofilm abundance in solely LB. Both genomes from these isolated lacked *pgaABCD* genes implicated in biofilm formation in *E. coli* (Wang et al. 2004). In addition, no biofilm formation capacity was observed in N27_F2 (ST5295) and N27_F3 (ST5295) when grown in LB. These genomes lacked the *bcsAB* genes associated with cellulose production during biofilm formation (Omadjela et al. 2013).

In contrast, E229_F1 (ST162), E215_F1 (ST5628), and E215_F3 (ST5628) showed moderate levels of biofilm formation capacity despite the lack of *pgaABCD* or *bcsAB* genes when grown in LB. However, trends became even more apparent when grown in LB with 3.5% NaCl, isolates lacking *pgaABCD*, *bcsABCD*, *fimA*, or *dgcT* did not show biofilm formation capacity. In contrast, isolates S845_F2 (ST144), N119_F1 (ST46), and S845_F3 (ST1434) showed no biofilm formation capacity despite the presence of all genes associated with biofilm production in both media. Biofilm formation capacity in the isolates likely depends on multiple factors such as pH, nutrients, and temperature, which can influence expression levels and up- or downregulating mutations within the genes (Dhanasekaran and Thajuddin 2016). Thus, while some patterns could be identified, analysis of the presence/absence of biofilm-associated genes is not sufficient to draw accurate conclusions about the phenotypes expressed.

The environment from which these isolates were sampled is characterized by water with a dynamic saline gradient (Lepranta and Myrberg 2009), which would create positive selection for halotolerant *E. coli*. The genomes from most isolates carried all genes associated with tolerance of osmotic stress (Figure S11) and all isolates grew in LB media with 3.5% NaCl, although generation time increased for almost all the isolates compared to the growth in LB media alone (Wilcoxon signed-rank exact test P -value = $3.012e-09$; Fig. 5A). Notably, the phylogroup E isolate S1008_F1(ST8979) grew particularly slow in the presence of 3.5% NaCl, with a generation time of 1.5 h, compared to generation times of 0.65–0.89 h for other isolates. In contrast, the isolate that did not form biofilm, N2036_F1(ST155), had an increased growth rate in LB with 3.5% NaCl (Fig. 5B).

Although significant differences in generation time between the phylogroups were not observed when grown in LB media (one-way ANOVA, P -value = .3) (Fig. 5C), generation time in 3.5% NaCl was dependent on phylogroup (one-way ANOVA, P -value = .021; Fig. 5D), with phylogroup B1 having a shorter generation time than phylogroup B2 (Tukey HSD, P -value = .019). No significant differences between phylogroups A and B1 (Tukey HSD, P -value = .92) and A and B2 (Tukey HSD, P -value = .087) were observed. The increased halotolerance in phylogroup B1 compared to B2 could be explained by the fact that phylogroup B1 is associated with naturalized *E. coli* strains, while B2 is more associated with animals and human sources (Clermont et al. 2013, Martak et al. 2020). This supports the hypothesis that *E. coli* in these marine environments could originate from various sources and also supports the use of additional approaches than determining concentrations of *E. coli* for assessment of fecal, or other sources of contamination.

Conclusion

While the original source of *E. coli* in this marine environment remains largely unclear, this study shows that coastal marine sed-

iments can harbour a variety of viable *E. coli*. The geographic distribution and genetic typing of these isolates was not random. This suggests that *E. coli* in this study might originate from several sources such as the WWTP or animals but may survive in the marine environment. All isolates were halotolerant and some formed biofilm supporting the hypothesis that the marine sediments select for *E. coli* with increased survival properties and that marine sediments, could be a potential reservoir for naturalized *E. coli*. This makes *E. coli* a questionable indicator for recent fecal contamination of marine sediments. Additionally, isolated *E. coli* harboured a variety of different ARGs and genes encoding virulence factors. This suggests there is a potential risk to human health, both through contact with potentially pathogenic STs and that marine sediments can serve as reservoirs for ARGs and genes for virulence factors.

Acknowledgments

The authors would like to acknowledge Clinical Genomics Lund, SciLifeLab, and Center for Translational Genomics (CTG), Lund University, for providing expertise and service with sequencing and analysis. Bioinformatic analyses were enabled by resources in the projects SNIC 2021/22–833 and SNIC 2021/23–604 provided by the Swedish National Infrastructure for Computing (SNIC) at UPPMAX (Uppsala Multidisciplinary Centre for Advanced Computational Science), partially funded by the Swedish Research Council through grant agreement number 2018–05973. The authors also acknowledge the City of Helsingborg, the crew of Sabella and Tage Rosenqvist for assistance with sampling.

Author contributions

Isabel K. Erb (Formal analysis, Investigation, Methodology, Visualization, Writing—original draft), Carolina Suarez (Conceptualization, Formal analysis, Methodology, Supervision, Visualization, Writing—review & editing), Ellinor M. Frank (Investigation, Visualization), Johan Bengtsson-Palme (Methodology, Validation), Elisabet Lindberg (Investigation, Resources), and Catherine J. Paul (Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing—review & editing)

Supplementary data

Supplementary data is available at *FEMSMC Journal* online.

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

Funding

Funding was provided by Lund University and Sweden Water Research AB as part of the Urban Bathing (Urbana Bad) project. J.B.P. acknowledges funding from the Data-Driven Life Science (DDL) program supported by the Knut and Alice Wallenberg Foundation (KAW 2020.0239), the Swedish Research Council (VR; grant 2019-00299) under the frame of JPI AMR (EMBARK; JPIAMR2019-109), and the Swedish Foundation for Strategic Research (FFL21-0174).

Data availability

Whole genome sequences are available at the NCBI Bioproject PRJNA967680. Codes for bioinformatic analysis will be made available upon request.

Reference

- Alcock BP, Huynh W, Chalil R et al. CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Res* 2023;**51**:D690–9.
- Alcock BP, Raphenya AR, Lau TTY et al. CARD 2020: antibiotic resistance surveillance with the Comprehensive Antibiotic Resistance Database. *Nucleic Acids Res* 2019;**48**:gkz935.
- Alves MS, Pereira A, Araújo SM et al. Seawater is a reservoir of multi-resistant *Escherichia coli*, including strains hosting plasmid-mediated quinolones resistance and extended-spectrum beta-lactamases genes. *Front Microbiol* 2014;**5**:426.
- Anastasi EM, Matthews B, Stratton HM et al. Pathogenic *Escherichia coli* found in sewage treatment plants and environmental waters. *Appl Environ Microbiol* 2012;**78**:5536–41.
- Balière C, Rincé A, Blanco J et al. Prevalence and characterization of shiga toxin-producing and enteropathogenic *Escherichia coli* in shellfish-harvesting areas and their watersheds. *Front Microbiol* 2015;**6**. <https://doi.org/10.3389/fmicb.2015.01356>.
- Beghain J, Bridier-Nahmias A, Le Nagard H et al. ClermonTyping: an easy-to-use and accurate in silico method for *Escherichia* genus strain phylotyping. *Microb Genom* 2018;**4**. <https://doi.org/10.1099/mgen.0.000192>.
- Berthe T, Ratajczak M, Clermont O et al. Evidence for coexistence of distinct *Escherichia coli* populations in various aquatic environments and their survival in estuary water. *Appl Environ Microbiol* 2013;**79**:4684–93.
- Buchfink B, Reuter K, Drost H-G. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat Methods* 2021;**18**:366–8.
- Can F, Azap OK, Seref C et al. Emerging *Escherichia coli* O25b/ST131 clone predicts treatment failure in urinary tract infections. *Clin Infect Dis* 2015;**60**:523–7.
- Cantalapiedra CP, Hernández-Plaza A, Letunic I et al. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol Biol Evol* 2021;**38**:5825–9.
- Carlos C, Pires MM, Stoppe NC et al. *Escherichia coli* phylogenetic group determination and its application in the identification of the major animal source of fecal contamination. *BMC Microbiol* 2010;**10**:161.
- Chaumeil P-A, Mussig AJ, Hugenholtz P et al. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* 2019;**36**:btz848.
- Clermont O, Christenson JK, Denamur E et al. The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups: a new *E. coli* phylo-typing method. *Environ Microbiol Rep* 2013;**5**:58–65.
- Clermont O, Condamine B, Dion S et al. The *E. coli* phylogroup of *Escherichia coli* is highly diverse and mimics the whole *E. coli* species population structure. *Environ Microbiol* 2021;**23**:7139–51.
- Coffey BM, Anderson GG. Biofilm formation in the 96-well microtiter plate. In: Filloux A, Ramos J-L (eds.), *Pseudomonas Methods and Protocols*. Vol. 1149. New York, NY: Springer, 2014, 631–41.
- Dang H, Lovell CR. Microbial surface colonization and biofilm development in marine environments. *Microbiol Mol Biol Rev* 2016;**80**:91–138.
- Denamur E, Clermont O, Bonacorsi S et al. The population genetics of pathogenic *Escherichia coli*. *Nat Rev Microbiol* 2021;**19**:37–54.
- Dhanasekaran D, Thajuddin N. *Microbial Biofilms*. Rijeka: IntechOpen, 2016.
- Dusek N, Hewitt AJ, Schmidt KN et al. Landscape-scale factors affecting the prevalence of *Escherichia coli* in surface soil include land cover type, edge interactions, and soil pH. *Appl Environ Microbiol* 2018;**84**:e02714–17.
- Finn TJ, Scriver L, Lam L et al. A comprehensive account of *Escherichia coli* sequence type 131 in wastewater reveals an abundance of fluoroquinolone-resistant clade A strains. *Appl Environ Microbiol* 2020;**86**:e01913–19.
- Flemming H-C, Wuertz S. Bacteria and archaea on Earth and their abundance in biofilms. *Nat Rev Microbiol* 2019;**17**:247–60.
- Frank EM, Ahlinder J, Jephson T et al. Marine sediments are identified as an environmental reservoir for *Escherichia coli*: comparing signature-based and novel amplicon sequencing approaches for microbial source tracking. *Sci Total Environ* 2024;**907**:167865.
- Gerba CP, McLeod JS. Effect of sediments on the survival of *Escherichia coli* in marine waters. *Appl Environ Microbiol* 1976;**32**:114–20.
- Grunwald JL, Rabinowitz P, Weissman SJ et al. Diversity of *Escherichia coli* found in the Salish Sea. *Front Mar Sci* 2022;**9**. <https://doi.org/10.3389/fmars.2022.967435>.
- Haile RW, Witte JS, Gold M et al. The health effects of swimming in ocean water contaminated by storm drain runoff. *Epidemiology* 1999;**10**:355–63.
- Halliday E, Gast RJ. Bacteria in beach sands: an emerging challenge in protecting coastal water quality and bather health. *Environ Sci Technol* 2011;**45**:370–9.
- Higgins J, Hohn C, Hornor S et al. Genotyping of *Escherichia coli* from environmental and animal samples. *J Microbiol Methods* 2007;**70**:227–35.
- Huerta-Cepas J, Szklarczyk D, Heller D et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* 2019;**47**:D309–14.
- Hyatt D, Chen G-L, LoCascio PF et al. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinf* 2010;**11**:119.
- Ishii S, Hansen DL, Hicks RE et al. Beach sand and sediments are temporal sinks and sources of *Escherichia coli* in Lake Superior. *Environ Sci Technol* 2007;**41**:2203–9.
- Jain C, Rodriguez-R LM, Phillippy AM et al. High throughput ANI analysis of 90 K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 2018;**9**:5114.
- Joensen KG, Scheutz F, Lund O et al. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. *J Clin Microbiol* 2014;**52**:1501–10.
- Johnson JR, Johnston BD, Delavari P et al. Phylogenetic backgrounds and virulence-associated traits of *Escherichia coli* isolates from surface waters and diverse animals in Minnesota and Wisconsin. *Appl Environ Microbiol* 2017;**83**:e01329–17.
- Johnson JR, Russo TA. Extraintestinal pathogenic *Escherichia coli*: “the other bad *E. coli*”. *J Lab Clin Med* 2002;**139**:155–62.
- Julian TR, Islam MA, Pickering AJ et al. Genotypic and phenotypic characterization of *Escherichia coli* isolates from feces, hands, and soils in rural Bangladesh via the Colilert Quanti-Tray System. *Appl Environ Microbiol* 2015;**81**:1735–43.
- Kaper JB, Nataro JP, Mobley HLT. Pathogenic *Escherichia coli*. *Nat Rev Microbiol* 2004;**2**:123–40.
- Kindle P, Zurfluh K, Nüesch-Inderbinen M et al. Phenotypic and genotypic characteristics of *Escherichia coli* with non-susceptibility to

- quinolones isolated from environmental samples on pig farms. *Porc Health Manag* 2019;**5**:9.
- Larsen MV, Cosentino S, Rasmussen S et al. Multilocus sequence typing of total-genome-sequenced bacteria. *J Clin Microbiol* 2012;**50**:1355–61.
- Lepranta M, Myrberg K. *Physical Oceanography of the Baltic Sea*. New York, NY: Springer, 2009.
- Low AJ, Koziol AG, Manninger PA et al. ConFindr: rapid detection of intraspecies and cross-species contamination in bacterial whole-genome sequence data. *PeerJ* 2019;**7**:e6995.
- Luo C, Walk ST, Gordon DM et al. Genome sequencing of environmental *Escherichia coli* expands understanding of the ecology and speciation of the model bacterial species. *Proc Natl Acad Sci USA* 2011;**108**:7200–5.
- Luo Y, Mao D, Rysz M et al. Trends in antibiotic resistance genes occurrence in the Haihe River, China. *Environ Sci Technol* 2010;**44**:7220–5.
- Madoshi BP, Kudirkiene E, Mtambo MMA et al. Characterisation of commensal *Escherichia coli* isolated from apparently healthy cattle and their attendants in Tanzania. *PLoS One* 2016;**11**:e0168160.
- Malberg Tetzschner AM, Johnson JR, Johnston BD et al. In silico genotyping of *Escherichia coli* isolates for extraintestinal virulence genes by use of whole-genome sequencing data. *J Clin Microbiol* 2020;**58**:e01269–20.
- Martak D, Henriot CP, Broussier M et al. High prevalence of human-associated *Escherichia coli* in wetlands located in eastern France. *Front Microbiol* 2020;**11**:552566.
- Martínez JA, Soto S, Fabrega A et al. Relationship of phylogenetic background, biofilm production, and time to detection of growth in blood culture vials with clinical variables and prognosis associated with *Escherichia coli* bacteremia. *J Clin Microbiol* 2006;**44**:1468–74.
- McCarthy DT, Jovanovic D, Lintern A et al. Source tracking using microbial community fingerprints: method comparison with hydrodynamic modelling. *Water Res* 2017;**109**:253–65.
- McGinnis SM, Burch T, Murphy HM. Assessing the risk of acute gastrointestinal illness (AGI) acquired through recreational exposure to combined sewer overflow-impacted waters in Philadelphia: a quantitative microbial risk assessment. *Microbial Risk Anal* 2022;**20**:100189.
- Nguyen F, Starosta AL, Arenz S et al. Tetracycline antibiotics and resistance mechanisms. *Biol Chem* 2014;**395**:559–75.
- Nicolas-Chanoine M-H, Bertrand X, Madec J-Y. *Escherichia coli* ST131, an intriguing clonal group. *Clin Microbiol Rev* 2014;**27**:543–74.
- Odonkor ST, Ampofo JK. *Escherichia coli* as an indicator of bacteriological quality of water: an overview. *Microbiol Res* 2013;**4**:2.
- Olowe OA, Adefoye OJ, Ajayeoba TA et al. Phylogenetic grouping and biofilm formation of multidrug resistant *Escherichia coli* isolates from humans, animals and food products in South-West Nigeria. *Sci Afr* 2019;**6**:e00158.
- Omadjela O, Narahari A, Strumillo J et al. BcsA and BcsB form the catalytically active core of bacterial cellulose synthase sufficient for in vitro cellulose synthesis. *Proc Natl Acad Sci USA* 2013;**110**:17856–61.
- Osińska A, Korzeniewska E, Harnisz M et al. Small-scale wastewater treatment plants as a source of the dissemination of antibiotic resistance genes in the aquatic environment. *J Hazard Mater* 2020;**381**:121221.
- Pachepsky YA, Shelton DR. *Escherichia coli* and fecal coliforms in freshwater and estuarine sediments. *Crit Rev Environ Sci Technol* 2011;**41**:1067–110.
- Parks DH, Chuvochina M, Rinke C et al. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res* 2022;**50**:D785–94.
- Parks DH, Imelfort M, Skennerton CT et al. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 2015;**25**:1043–55.
- Picard B, Garcia JS, Gouriou S et al. The link between phylogeny and virulence in *Escherichia coli* extraintestinal infection. *Infect Immun* 1999;**67**:546–53.
- QGIS Development Team. *QGIS Geographic Information System*. Beaverton, OR: Open Source Geospatial Foundation Project, 2023.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing, 2021.
- Raboni M, Gavasci R, Torretta V. Assessment of the fate of *Escherichia coli* in different stages of wastewater treatment plants. *Water Air Soil Pollut* 2016;**227**:455.
- Redgrave LS, Sutton SB, Webber MA et al. Fluoroquinolone resistance: mechanisms, impact on bacteria, and role in evolutionary success. *Trends Microbiol* 2014;**22**:438–45.
- Riley LW. Pandemic lineages of extraintestinal pathogenic *Escherichia coli*. *Clin Microbiol Infect* 2014;**20**:380–90.
- Ross BN, Rojas-Lopez M, Cieza RJ et al. The role of long polar fimbriae in *Escherichia coli* O104:H4 adhesion and colonization. *PLoS One* 2015;**10**:e0141845.
- Rumball NA, Mayer HC, McLellan SL. Selective survival of *Escherichia coli* phylotypes in freshwater beach sand. *Appl Environ Microbiol* 2021;**87**:e02473–20.
- Salam S, McDaniel R, Bleakley B et al. Variability of *E. coli* in streambed sediment and its implication for sediment sampling. *J Contam Hydrol* 2021;**242**:103859.
- Sarowska J, Futoma-Koloch B, Jama-Kmieciak A et al. Virulence factors, prevalence and potential transmission of extraintestinal pathogenic *Escherichia coli* isolated from different sources: recent reports. *Gut Pathog* 2019;**11**:10.
- Schang C, Lintern A, Cook PLM et al. *Escherichia coli* survival and transfer in estuarine bed sediments: *E. coli* survival and transfer in estuarine bed sediments. *River Res Apps* 2018;**34**:606–14.
- Shaw KJ, Rather PN, Hare RS et al. Molecular genetics of aminoglycoside resistance genes and familial relationships of the aminoglycoside-modifying enzymes. *Microbiol Rev* 1993;**57**:138–63.
- Solo-Gabriele HM, Wolfert MA, Desmarais TR et al. Sources of *Escherichia coli* in a coastal subtropical environment. *Appl Environ Microbiol* 2000;**66**:230–7.
- Souvorov A, Agarwala R, Lipman DJ. SKESA: strategic k-mer extension for scrupulous assemblies. *Genome Biol* 2018;**19**:153.
- Sprouffske K, Wagner A. Growthcurver: an R package for obtaining interpretable metrics from microbial growth curves. *BMC Bioinf* 2016;**17**:172.
- Spurbeck RR, Dinh PC, Walk ST et al. *Escherichia coli* isolates that carry vat, fyuA, chuA, and yfcV efficiently colonize the urinary tract. *Infect Immun* 2012;**80**:4115–22.
- Tenaillon O, Skurnik D, Picard B et al. The population genetics of commensal *Escherichia coli*. *Nat Rev Microbiol* 2010;**8**:207–17.
- Touchon M, Perrin A, de Sousa JAM et al. Phylogenetic background and habitat drive the genetic diversification of *Escherichia coli*. *PLoS Genet* 2020;**16**:e1008866.
- Tropea E, Hynds P, McDermott K et al. Environmental adaptation of *E. coli* within private groundwater sources in southeastern Ontario: implications for groundwater quality monitoring and human health. *Environ Pollut* 2021;**285**:117263.
- van der Woude MW, Henderson IR. Regulation and function of Ag43 (Flu). *Annu Rev Microbiol* 2008;**62**:153–69.

- van Elsas JD, Semenov AV, Costa R et al. Survival of *Escherichia coli* in the environment: fundamental and public health aspects. *ISME J* 2011;**5**:173–83.
- Vignaroli C, Luna GM, Pasquaroli S et al. Epidemic *Escherichia coli* ST131 and *Enterococcus faecium* ST17 in coastal marine sediments from an Italian beach. *Environ Sci Technol* 2013;**47**:13772–80.
- Walk ST, Alm EW, Calhoun LM et al. Genetic diversity and population structure of *Escherichia coli* isolated from freshwater beaches. *Environ Microbiol* 2007;**9**:2274–88.
- Wang H, Hou L, Liu Y et al. Horizontal and vertical gene transfer drive sediment antibiotic resistome in an urban lagoon system. *J Environ Sci* 2021;**102**:11–23.
- Wang X, Preston JF, Romeo T. The *pgaABCD* locus of *Escherichia coli* promotes the synthesis of a polysaccharide adhesin required for biofilm formation. *J Bacteriol* 2004;**186**:2724–34.
- Whitman RL, Harwood VJ, Edge TA et al. Microbes in beach sands: integrating environment, ecology and public health. *Rev Environ Sci Biotechnol* 2014;**13**:329–68.
- Wirth T, Falush D, Lan R et al. Sex and virulence in *Escherichia coli*: an evolutionary perspective. *Mol Microbiol* 2006;**60**:1136–51.
- Wright ME, Solo-Gabriele HM, Elmir S et al. Microbial load from animal feces at a recreational beach. *Mar Pollut Bull* 2009;**58**:1649–56.
- Wu C, Zhang G, Xu W et al. New estimation of antibiotic resistance genes in sediment along the Haihe River and Bohai Bay in China: a comparison between single and successive DNA extraction methods. *Front Microbiol* 2021;**12**:705724.
- Yu D, Ryu K, Otto SJG et al. Differential survival of potentially pathogenic, septicemia- and meningitis-causing *E. coli* across the wastewater treatment train. *npj Clean Water* 2022;**5**:33.
- Zhi S, Stothard P, Banting G et al. Characterization of water treatment-resistant and multidrug-resistant urinary pathogenic *Escherichia coli* in treated wastewater. *Water Res* 2020;**182**:115827.
- Zhou M, Ding X, Ma F et al. Long polar fimbriae contribute to pathogenic *Escherichia coli* infection to host cells. *Appl Microbiol Biotechnol* 2019;**103**:7317–24.
- Zhou M, Yang Y, Wu M et al. Role of long polar fimbriae type 1 and 2 in pathogenesis of mammary pathogenic *Escherichia coli*. *J Dairy Sci* 2021;**104**:8243–55.
- Zhou Z, Alikhan N-F, Mohamed K et al. The Enterobase user's guide, with case studies on *Salmonella* transmissions, *Yersinia pestis* phylogeny, and *Escherichia* core genomic diversity. *Genome Res* 2020;**30**:138–52.
- Zieliński W, Korzeniewska E, Harnisz M et al. Wastewater treatment plants as a reservoir of integrase and antibiotic resistance genes—an epidemiological threat to workers and environment. *Environ Int* 2021;**156**:106641.