Extended-spectrum beta-lactamase-producing Escherichia coli and antimicrobial resistance in municipal and hospital wastewaters in Czech Republic

culture-based and metagenomic approaches

Kutilova, Iva; Medvecky, Matej; Leekitcharoenphon, Pimlapas; Munk, Patrick; Masarikova, Martina; Davidova-Gerzova, Lenka; Jamborova, Ivana; Bortolaia, Valeria; Pamp, Sünje Johanna; Dolejska, Monika

Published in:
Environmental Research

Link to article, DOI:
10.1016/j.envres.2020.110487

Publication date:
2021

Document Version
Peer reviewed version

Journal Pre-proof

Extended-spectrum beta-lactamase-producing *Escherichia coli* and antimicrobial resistance in municipal and hospital wastewaters in Czech Republic: culture-based and metagenomic approaches

Iva Kutilova, Matej Medvecky, Pimlapas Leekitcharoenphon, Patrick Munk, Martina Masarikova, Lenka Davidova-Gerzova, Ivana Jamborova, Valeria Bortolaia, Sünje J. Pamp, Monika Dolejska

PII: S0013-9351(20)31384-0

DOI: https://doi.org/10.1016/j.envres.2020.110487

Reference: YENRS 110487

To appear in: *Environmental Research*

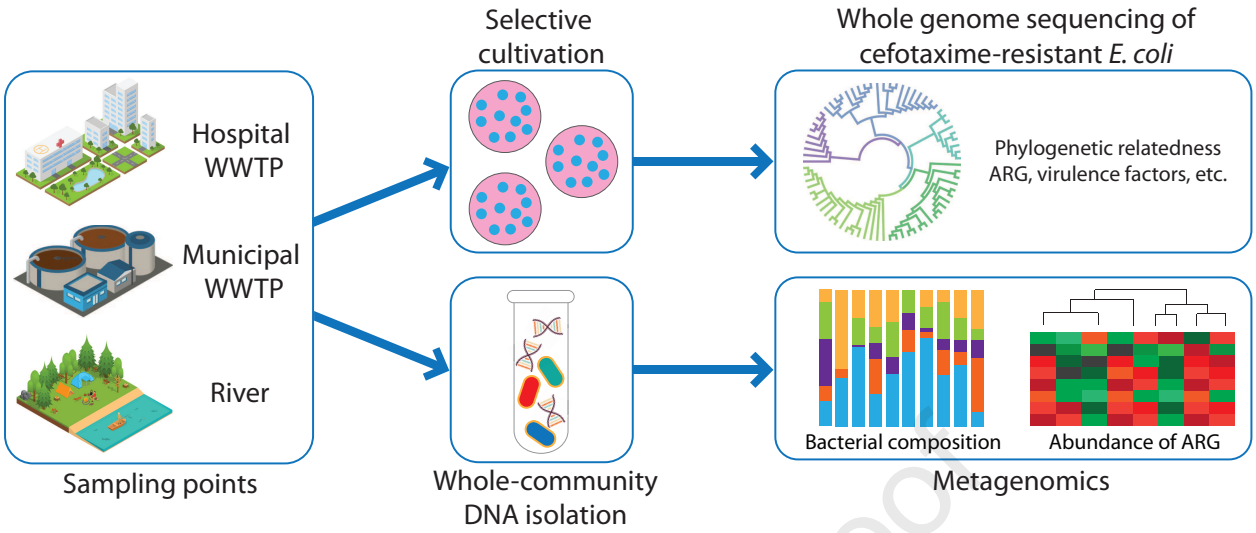
Received Date: 1 June 2020

Revised Date: 31 October 2020

Accepted Date: 12 November 2020

Please cite this article as: Kutilova, I., Medvecky, M., Leekitcharoenphon, P., Munk, P., Masarikova, M., Davidova-Gerzova, L., Jamborova, I., Bortolaia, V., Pamp, S.J., Dolejska, M., Extended-spectrum beta-lactamase-producing *Escherichia coli* and antimicrobial resistance in municipal and hospital wastewaters in Czech Republic: culture-based and metagenomic approaches, *Environmental Research*, https://doi.org/10.1016/j.envres.2020.110487.

Iva Kutilova: writing, investigation, visualization, formal analysis, funding acquisition. **Matej Medvecky:** investigation, visualization, formal analysis. **Pimlapas Leekitcharoenphon**: supervision. **Patrick Munk**: investigation, formal analysis. **Martina Masarikova**: investigation. **Lenka Davidova-Gerzova**: investigation. **Ivana Jamborova**: investigation. **Valeria Bortolaia**: supervision. **Sünje J. Pamp**: supervision. **Monika Dolejska**: supervision, conceptualization, formal analysis, funding acquisition.



Extended-spectrum beta-lactamase-producing *Escherichia coli* and antimicrobial resistance in municipal and hospital wastewaters in Czech Republic: culture-based and metagenomic approaches

Iva Kutilova[1,2], Matej Medvecky[1,3], Pimlapas Leekitcharoenphon[4], Patrick Munk[4], Martina Masarikova[1,5], Lenka Davidova-Gerzova[1,5], Ivana Jamborova[1], Valeria Bortolaia[4], Sünje J. Pamp[4] and Monika Dolejska[1,2,3*]

[1]CEITEC VFU, University of Veterinary and Pharmaceutical Sciences Brno, Palackeho tr. 1946/1, 612 42 Brno, Czech Republic

[2]Department of Biology and Wildlife Diseases, Faculty of Veterinary Hygiene and Ecology, University of Veterinary and Pharmaceutical Sciences Brno, Palackeho tr. 1946/1, 612 42 Brno, Czech Republic

[3]Biomedical Center, Faculty of Medicine, Charles University, Plzen, Czech Republic

[4]Research Group for Genomic Epidemiology, National Food Institute, Technical University of Denmark, Kgs. Lyngby 2800, Denmark

[5]Department of Infectious Diseases and Microbiology, Faculty of Veterinary Medicine, University of Veterinary and Pharmaceutical Sciences Brno, Palackeho tr. 1946/1, 612 42 Brno, Czech Republic

[*]Corresponding author. CEITEC VFU, University of Veterinary and Pharmaceutical Sciences Brno, Palackeho tr. 1946/1, 612 42 Brno, Czech Republic; E-mail: monika.dolejska@gmail.com

Abstract

Wastewaters serve as important hot spots for antimicrobial resistance and monitoring can be used to analyse the abundance and diversity of antimicrobial resistance genes at the level of large bacterial and human populations. In this study, whole genome sequencing of beta-lactamase-producing *Escherichia coli* and metagenomic analysis of whole-community DNA were used to characterize the occurrence of antimicrobial resistance in hospital, municipal and river waters in the city of Brno (Czech Republic).

Cefotaxime-resistant *E. coli* were mainly extended-spectrum beta-lactamase (ESBL) producers (95.6%, n=158), of which the majority carried $bla_{\text{CTX-M}}$ (98.7%; n=151) and were detected in all water samples except the outflow from hospital wastewater treatment plant. A wide phylogenetic diversity was observed among the sequenced *E. coli* (n=78) based on the detection of 40 sequence types and single nucleotide polymorphisms (average number 34,666 ± 15,710) between strains. The metagenomic analysis revealed a high occurrence of bacterial genera with potentially pathogenic members, including *Pseudomonas*, *Escherichia*, *Klebsiella*, *Aeromonas*, *Enterobacter* and *Arcobacter* (relative abundance > 50%) in untreated hospital and municipal wastewaters and predominance of environmental bacteria in treated and river waters. Genes encoding resistance to aminoglycosides, beta-lactams, quinolones and macrolides were frequently detected, however $bla_{\text{CTX-M}}$ was not found in this dataset which may be affected by insufficient sequencing depth of the samples.

The study pointed out municipal treated wastewater as a possible source of multi-drug resistant *E. coli* and antimicrobial resistance genes for surface waters. Moreover, the combination of two different approaches provided a more holistic view on antimicrobial resistance in water environments. The culture-based approach facilitated insight into the dynamics of ESBL-producing *E. coli* and the metagenomics shows abundance and diversity of bacteria and antimicrobial resistance genes vary across water sites.

Keywords: *E. coli*; wastewater treatment plants; ESBL; whole genome sequencing

Funding

This work was supported by Central European Institute of Technology (CEITEC) 2020 from the Czech Ministry of Education, Youth and Sports within the National Programme for Sustainability II. **[**grant number LQ1601]; the Ministry of Health of the Czech Republic [grant number NV18-09-00605]; Internal Grant Agency of the Veterinary and Pharmaceutical University Brno [grant numbers 204/2018/FVHE, 205/2019/FVHE]. The contributions by the authors from the Technical University of Denmark were partly supported by the European Union's Framework program for Research and Innovation, Horizon 2020 [grant number 643476]; the Novo Nordisk Foundation [grant number NNF16OC0021856: Global Surveillance of Antimicrobial Resistance].

1. Introduction

The availability of drinking water, appropriate sewage system and treatment of wastewaters (WWs) are fundamental to the maintenance and further improvement of public health (Bürgmann et al., 2018). Wastewater treatment plants (WWTPs) represent places where sewage (containing bacteria and other microorganisms, chemical residues and nutrients) from different sources such as hospitals, communities, industry and agriculture is interacting (Amos et al., 2014). Conventional WWTPs have capacity to reduce bacteria only by 10-100 fold, which allows bacterial spread via treated water (Karkman et al., 2018). Therefore, WWTPs are considered hot spots for spread of antimicrobial-resistant bacteria (ARB) and antimicrobial resistance genes (ARG) in the environment and were also pointed out to facilitate favourable conditions for horizontal gene transfer during the treatment processes (Amos et al., 2014; Garner et al., 2018). The degree to which antimicrobials select for ARB in WWTPs remains to be elucidated and environmental studies taking into account the interaction of whole community of bacteria with ecological factors are needed (Manaia et al., 2018).

To study antimicrobial resistance (AMR) in WWs, culture-based and culture-independent (especially quantitative PCR and metagenomics) methods are used. These methods give different types of information and a combination of both approaches is recommended (Manaia et al., 2018).

Metagenomics is based on high throughout sequencing of all DNA extracted from sample (e.g. sewage, animal faeces, etc.) and the sufficient sequencing depth of a sample is crucial. One million and at least 80 million reads per sample have been suggested as minimal values to be accurate within 1% of true beta-diversity and to observe a comprehensive composition of antimicrobial resistance gene families, respectively (Gweon et al., 2019). Though in the same study, it was also demonstrated that even 200 million reads were not sufficient to

establish allelic diversity of ARG in the sample from WWTP effluent. The global sewage project relies extensively on metagenomic studies and in a recent study over 1.4 TB of sequencing data from untreated sewage were analysed and differences in the abundance and diversity of ARG between continents were observed (Hendriksen et al., 2019). The most abundant ARG included those conferring resistance to macrolides, beta-lactams, tetracycline and aminoglycosides in all locations. The authors suggested metagenomic analysis of sewage as a suitable approach for global AMR monitoring. Nevertheless, the application of metagenomic analyses should be considered wisely according to the aim of the study. Huge changes in bacterial composition during the wastewater treatment that can influence the abundance of ARG has been demonstrated (Bengtsson-Palme et al. 2016). Therefore, it is not recommended to use only metagenomic analyses for comparison of AMR and prediction of antimicrobial selective pressure between untreated and treated water and information from culture-based methods should be included too.

Escherichia coli is known as a human commensal strain as well as a human pathogen capable of causing intestinal (e.g. gastroenteritis) and extraintestinal (e.g. bloodstream and urinary tract) infections. Certain groups of virulence factors and similar phylogenetic backgrounds are characteristic of extraintestinal pathogenic *E. coli* (ExPEC) which are highly relevant in the clinic also due to AMR occurrence (Russo and Johnson, 2003). Particularly, *E. coli* resistant to cefotaxime (CTX) and other medically important cephalosporins by production of extended-spectrum beta-lactamases (ESBLs) pose a serious clinical threat (Paulshus et al., 2019). Hospitals together with WWTPs are considered as the main source of ESBL-producing *E. coli* for the environment (Gomi et al., 2017a). Certain *E. coli* sequence types (STs) including ST69, ST73, ST95 and ST131 are frequently associated with extraintestinal diseases and show a wide variety in encoded virulence factors and AMR determinants (Dale and Woodford, 2015). Especially pandemic ExPEC ST131, resistant to cefotaxime (by CTX-M-15

ESBL) and/or fluoroquinolones, is involved in millions of infections per year (Pitout and Devinney, 2017).

The aim of this study was to characterize the occurrence and spread of AMR in hospital, municipal and river waters in the city of Brno (Czech Republic). Three different methods were used to fulfil this aim: (1) enumeration of total vs CTX-resistant *E. coli*, (2) phenotyping and genotyping of representative CTX-resistant *E. coli* isolates, and (3) metagenomic analysis of whole-community DNA of each water sample.

2. Materials and methods

2.1. Sampling sites

The water sampling was conducted during October 2016 in the city of Brno (approximately 400,000 inhabitants), Czech Republic. A total of six sewage samples were collected: three samples from the University hospital Brno Bohunice (49°10'49.4"N 16°34'20.7"E) were obtained on one day and two days later, two samples from the municipal wastewater treatment plant (WWTP) Brno Modrice (49°07'52.9"N 16°37'57.7"E) and one from the Svratka river (49°07'49.3"N 16°37'38.0"E) taken upstream of the outflow from municipal WWTP (Fig. 1). The hospital samples included inflow to the hospital WWTP (Hospital-WWTP-In), outflow from the hospital WWTP (Hospital-WWTP-Out) and raw hospital sewage (Hospital-Raw) flowing to municipal WWTP. From the municipal WWTP, inflow (Municipal-WWTP-In) and outflow (Municipal-WWTP-Out) were sampled. In the hospital WWTP, wastewater from infection clinics is treated and a chlorine disinfection is used as a last step of wastewater treatment. The municipal WWTP is based on two-stage treatment, i.e. mechanical and biological, with usage of anaerobic sludge stabilisation. From each sampling, a sample of 1 L

was collected in a sterile glass bottle. The samples were transported in a cooling box, kept at 4 °C and processed during the day of sampling.

2.2. E. coli *isolation*

The water samples were diluted by serial ten-fold dilution method to obtain countable plates or the samples were concentrated using filtering (0.22 µm; Sigma-Aldrich, US) before plating. For enumeration of *E. coli* and other coliform bacteria, chromogenic medium Brilliance™ *E. coli*/coliform Selective Agar (Oxoid, UK) with and without cefotaxime (CTX, 2 mg/L; Sigma-Aldrich, US) was used and plates were incubated at 37 °C overnight. All samples were tested in triplicates. The proportion of CTX-resistant *E. coli* and other coliform bacteria out of the total counts of *E. coli*/coliforms were calculated and Fisher's exact test (with p-value < 0.05 indicating significant difference) was used to evaluate influence of wastewater treatment process (Municipal-WWTP-In vs Municipal-WWTP-Out) on CTX-resistant bacteria selection.

Up to 40 colonies of presumptive CTX-resistant *E. coli* were collected from the different plates of each water sample trying to encompass as much diversity as possible. The selected colonies were subcultured on MacConkey agar (Oxoid, UK) and stored at -80 °C. Species identification was verified using MALDI-TOF MS (Microflex LT, Bruker Daltonics, Germany) and only *E. coli* isolates were subjected to further typing.

2.3. Phenotyping and genotyping of cefotaxime-resistant E. coli

The isolates were tested for the production of ESBL/AmpC enzymes by D68C1 AmpC & ESBL Detection Set (Mast Diagnostics, UK) followed by double-disk synergy test (CLSI 2008) in case of questionable results. The susceptibility to 15 antimicrobial compounds (Oxoid, UK) was determined by disk diffusion method (CLSI, 2015a). The tested

antimicrobials included amoxicillin-clavulanic acid (30 μg), ampicillin (10 μg), ceftazidime (30 μg), cephalothin (30 μg), chloramphenicol (30 μg), ciprofloxacin (5 μg), ertapenem (10 μg), gentamicin (10 μg), imipenem (10 μg), meropenem (10 μg), nalidixic acid (30 μg), streptomycin (10 μg), sulphamethoxazole-trimethoprim (23,75/1,25 μg), compound sulphonamides S3 (300 μg) and tetracycline (30 μg). The inhibition zones diameters were interpreted according to CLSI, 2017 standards except for cephalothin where CLSI, 2015b standard was used.

A heatmap with clustering based on antimicrobial susceptibility profiles of all isolates was constructed using gplots (Warnes et al., 2019) and heatmap.plus (Day, 2015) libraries from the R environment v3.5.1 and edited in Inkscape v0.92 (https://inkscape.org/). All isolates were tested for the presence of beta-lactamase genes (*bla*$_{\text{CTX-M}}$, *bla*$_{\text{GES}}$, *bla*$_{\text{OXA}}$, *bla*$_{\text{SHV}}$, *bla*$_{\text{TEM}}$) by polymerase chain reaction (PCR) (Dobiasova et al. 2013). Primers for detection of *bla*$_{\text{GES}}$ gene were designed for this study (GES_start 5'-ATGCGCTTCATTCACGC-3'; GES_end 5'-CTATTTGTCCGTGCTCAGGA-3') according to the sequence of *bla*$_{\text{GES-1}}$ (GenBank accession AF355189). AmpC producers were additionally screened for the following genes (*bla*$_{\text{ACC}}$, *bla*$_{\text{ACT}}$, *bla*$_{\text{BIL}}$, *bla*$_{\text{CMY}}$ *bla*$_{\text{DHA}}$, *bla*$_{\text{FOX}}$, *bla*$_{\text{LAT}}$ and *bla*$_{\text{MOX}}$) (Pérez-pérez and Hanson 2002) and PCR amplicons were sequenced using Sanger-based technology. *E. coli* ST131 genotype was tested by multiplex PCR (Johnson et al., 2009).

The clonality of all *E. coli* isolates was determined by pulsed-field gel electrophoresis (PFGE) (Centers for Disease Control and Prevention, 2004) using *Xba*I enzyme (Takara Bio Inc., Japan). Cluster analysis of macrorestriction patterns was performed using BioNumerics 6.6 software (Applied Maths, Belgium) based on Dice similarity indices. Isolates with ≥ 85% similarity of PFGE patterns were categorized into the same cluster (Carrico et al., 2005).

At least one isolate per sampling site from each PFGE cluster was selected for whole genome sequencing (WGS).

2.4. WGS, assembly and data analysis of CTX-resistant E. coli

The DNA from isolates was extracted by NucleoSpin Tissue kit (Macherey-Nagel, Germany), used for preparation of DNA libraries (Nextera XT DNA Library Preparation Kit, Illumina, Inc., USA) and sequenced on the NextSeq (2 x 150bp paired-end sequencing, NextSeq 500/550 High Output Kit v2, Ilumina) or the MiSeq (2 x 250 bp paired-end sequencing, MiSeq Reagents Kits v2, Illumina) platforms (Illumina). All steps were performed according to the manufacturer's instructions. Raw Illumina paired-end reads were quality trimmed by Trimmomatic v0.32 (Bolger et al., 2014) and assembled using the *de novo* assembler SPAdes v3.12.0 (Bankevich et al., 2012). The obtained contigs were analysed for the presence of ARG and chromosomal point mutations associated with resistance by ResFinder 3.1 (Zankari et al., 2012) and for plasmid replicons by PlasmidFinder 2.0 (Carattoli et al., 2014). The presence of chromosomal mutations was evaluated based on the criteria that one single chromosomal mutation in *gyrA* gene confers low-level resistance to quinolones, and several mutations in DNA gyrase genes (*gyrA* and *gyrB*) and topoisomerase IV genes (*parC* and *parE*) are required to increase the level of quinolones resistance in *Enterobacteriaceae* (Correia et al. 2017). ARG or plasmid replicons were considered present if length coverage and identity to the reference sequence were 100% and ≥ 95%, respectively. STs and phylogenetic group of *E. coli* isolates were determined by MLST 2.0 (Larsen et al., 2012) and ClermonTyping tools (Beghain et al., 2018), respectively. The unknown STs were proceeded by analysis using EnteroBase v1.1.2 (Zhou et al., 2019) and new sequence type (ST) numbers were assigned.

The virulence factor database (VFDB; on September 2019) (Liu et al., 2019) was used to identify ExPEC isolates based on the definition that ExPEC strains contain two or more of the following virulence factors: *papA* and/or *papC* (P fimbriae), *sfa/focDE* (S and F1C fimbriae), *afa/draBC* (Dr-binding adhesins), *kpsM II* (group 2 capsule) and *iutA* (aerobactin receptor)

(Peirano et al., 2013). The length coverage and identity of virulence gene to the reference sequence had to be 100% and ≥ 90%, respectively.

For phylogenetic analysis, quality-trimmed reads were mapped against *E. coli* str. K-12 substr. MG1655 reference genome (GenBank accession U00096) employing bowtie2 v2.3.4.2 (Langmead and Salzberg, 2012). Samtools v1.3.1 (Li et al., 2009) was used for the format conversions and generation of the mpileup files. Single-nucleotide polymorphisms (SNPs) were then detected by VarScan v2.4.3 (Koboldt et al., 2012) with a minimum read depth of 6 (general tree constructed from all the sequenced samples) or 8 (supplementary trees inferred from samples belonging to particular STs of interest), respectively; minimum base quality of 20; and variant allele frequency ≥ 0.8. Moreover, all the sites in which at least one sample had read depth < 6 or 8, respectively, were removed.

Filtered SNPs of all the datasets were concatenated and analysed separately by jModeltest v2.1.10 (Darriba et al., 2015) in order to find the best-fitting nucleotide substitution model. Under Akaike information criterion, GTR model with the Γ model of rate heterogeneity was estimated as the best-fitting for datasets of SNPs from all the sequenced samples, as well as datasets comprised of samples belonging to ST38 and ST635. Plain GTR model was estimated as the most suitable for datasets generated from ST10 and ST131 samples. Maximum-likelihood analyses were performed by RAxML v8.2.10 (Stamatakis, 2014), robustness of the inferred topologies was assessed by 500 bootstrap replicate analyses. Sample IM1c (Municipal-WWTP-In isolate 1c) was used as an outgroup in the analysis of the largest dataset. Resulting tree topologies were visualised via iTOL v4.3.2 (Letunic and Bork, 2019) and edited in Inkscape v0.92 (www.inkscape.org). SNP distances among samples were calculated employing Biostrings library (Pagès et al., 2019) from the R environment v3.5.1.

2.5. Sequencing of whole-community DNA and data analysis

The whole-community DNA from each water sample was isolated using phenol-chloroform for obtaining large amounts of DNA followed by PowerWater DNA Isolation kit (MO BIO Laboratories, Inc., USA) for removal of poor-quality DNA. The DNA was used for the library preparation (Nextera XT DNA Library Preparation Kit) followed by sequencing on Illumina MiSeq platform using MiSeq Reagents Kits v3 (2 x 250 bp paired-end sequencing). The reads were quality- and adaptor-trimmed (quality threshold Q ≥ 20 and minimum length of 50 bp) using bbduk2 which is part of the suite bbtools version 36.49 (https://sourceforge.net/projects/bbmap/).

The metagenomic dataset was analysed using MGmapper (MetaGenomics mapper) 2.4 (Petersen et al., 2017) based on mapping reads to reference sequences and which allows two modes: best mode (a read can be assigned to one reference sequence) and full mode (a read can be assigned to more databases). The databases are based on data from NCBI GenBank except databases managed by the Center for Genomic Epidemiology namely ResFinder, PlasmidFinder and VirulenceFinder (Joensen et al., 2014), and the database called waterborne *E. coli*-db which consists of genome assemblies of *E. coli* isolates (n=78) obtained within this study. The metagenomic analysis was run (on May 2018) in the best mode for the respective databases in the following order: waterborne *E. coli*-db (created on May 2018; not complete genomes), Bacteria (complete genomes), Bacteria_draft (not complete genomes), HumanMicrobiome, Archaea, Virus, Protozoa, Human, Vertebrate_mammals, Vertebrate_other and Invertebrates. Full mode was used for ResFinder (resistance genes, on November 2018), Plasmid (plasmid sequences from NCBI GenBank; on May 2018), PlasmidFinder (plasmid replicons for Enterobacteriaceae and Gram-positive bacteria; on May 2019) and VirulenceFinder (virulence factors, on November 2018) databases. The paired reads mapping was used for all above mentioned databases except PlasmidFinder where

single-end read mapping (requiring at least 50 bp alignment for any single read) were preferable due to the short length of plasmid replicon sequences.

The number of reads mapping to bacterial databases (waterborne *E. coli*-db, Bacteria, Bacteria_draft, HumanMicrobiome) was summarized per sample and bacterial abundances were calculated as percentage per sample (relative abundance).

ARG were clustered according to 95% similarity of the sequence (CD-HIT_EST). The relative abundance of ARG clusters and plasmid replicons was calculated as fragments per kilobase of reference per million mapped reads (FPKM). The FPKM values were log-transformed and visualised in heatmaps (for antimicrobial classes and ARG) as previously described (Hendriksen et al., 2019) using pheatmap (Kolde, 2018) and viridis (Garnier, 2018) for color scale from the R environment v3.5.1 and R studio v1.1.456 (R Core Team 2018) and edited in Inkscape v0.92 (https://inkscape.org/).

3. Results

3.1. Proportion of cefotaxime-resistant E. coli *and coliform bacteria among total* E. coli *and coliforms*

Presumptive *E. coli* and other coliform bacteria were enumerated by counting colonies on cultivation media with and without cefotaxime (Table S1). The proportion of CTX resistant *E. coli* and other coliform bacteria out of total *E. coli* and other coliform bacteria was 4.83% and 16.67% in Hospital-WWTP-In, 0.67% and 11.08% in Hospital-Raw, 0.25% and 0.08% in Municipal-WWTP-In, 0.87% and 0.3% in Municipal-WWTP-Out and 2.13% and 6.67% in river water respectively. No *E. coli* or coliform bacteria (CFU/10 ml) were detected in Hospital-WWTP-Out. The counts of total and CTX-resistant *E. coli*/other coliform bacteria were in range 134-746-fold lower in Municipal-WWTP-Out than in Municipal-WWTP-In. However, the proportions of CTX-resistant *E. coli*/other coliform bacteria among total *E.*

coli/other coliform bacteria were significantly higher (p-value < 0.01) in Municipal-WWTP-Out compared to Municipal-WWTP-In.

3.2. Beta-lactamases among CTX-resistant E. coli

A total of 158 CTX-resistant *E. coli* were obtained from the different samples (for the numbers of isolates according their origin see Table 1). No CTX-resistant isolates were obtained in Hospital-WWTP-Out. Most CTX-resistant *E. coli* showed ESBL-production (151/158, 95.6%) and only a minority of isolates produced AmpC beta-lactamases (9/158, 5.7%). Among these, two isolates produced both ESBL and AmpC beta-lactamases. Most ESBL-producing isolates were positive for $bla_{\text{CTX-M}}$ (149/151, 98.7%) while the AmpC phenotype was associated with bla_{CMY} (6/9, 66.7%) and bla_{DHA} (2/9, 22.2%). The occurrence of other beta-lactamase genes varied according to the origin of samples. Genes bla_{TEM} (87/158, 55.1%) and bla_{OXA} (74/158, 46.8%) were frequently detected, while bla_{SHV} (2/158, 1.3%) and bla_{GES} (1/158, 0.6%) was detected at low occurrence (Table 1, Fig. 2, Table S2).

3.3. Antimicrobial resistance phenotype of CTX-resistant E. coli

The isolates exhibited diverse resistance phenotypes and did not group together into the same cluster according to their origin (Fig. 2, Table S2). All *E. coli* isolates (n=158) showed resistance to ampicillin (100%, penicillins) and cephalotin (100%; first generation of cephalosporins) which is explained by their selective isolation on media with cefotaxime. In total, 131 isolates (82.9%; n=158) were multi-drug resistant (i.e. resistant to compounds from three or more antimicrobial groups). More than half of the isolates were resistant to ceftazidime (64.6%; third generation of cephalosporins), sulphamethoxazole-trimethoprim (63.3%), nalidixic acid (60.8%; quinolones), streptomycin (58.2%; aminoglycosides) and tetracycline (57.6%). Resistance to other tested antimicrobials was as follows: sulphonamides (47.5%), gentamicin (43.7%; aminoglycosides), amoxicillin-clavulanic acid (34.8%;

penicillins and beta-lactamase inhibitor), chloramphenicol (21.6%; amphenicols) and ciprofloxacin (21.5%; fluoroquinolones). No isolate was resistant to carbapenems.

3.4. Genetic diversity of *CTX-resistant* E. coli

E. coli isolates showed high genetic diversity and were divided into 66 clusters based on ≥ 85% similarity of PFGE patterns (Table S2). The isolates from Municipal-WWTP-In and Municipal-WWTP-Out were divided into 19 and 25 clusters, respectively, whereas in hospital wastewaters there were 14 clusters for Hospital-WWTP-In and only 10 clusters for Hospital-Raw. Overall, most clusters contained isolates of the same origin. Despite generally high genetic diversity, isolates with indistinguishable PFGE profiles were found in Hospital-Raw and Municipal-WWTP-In (n=2); and in Municipal-WWTP-In and Municipal-WWTP-Out (n=5). Eighteen isolates (18/158, 11.4%) were classified as ST131 and originated from Hospital-Raw (3/39, 7.7%), Municipal-WWTP-In (10/38, 26.3%) and Municipal-WWTP-Out (5/35, 14.3%) (Table S2).

3.5. WGS of CTX-resistant E. coli – *MLST and phylogenetic group typing*

Based on the phenotyping and genotyping and the origin of the isolates, 78 representative *E. coli* isolates were subjected to WGS: 13 isolates from Hospital-Raw, 13 isolates from Hospital-WWTP-In, 20 isolates from Municipal-WWTP-In, 28 isolates from Municipal-WWTP-Out, and 4 isolates from river (Table S2). Among the sequenced *E. coli* isolates, 40 STs were identified. *E. coli* ST38 (12.8%), ST131 (10.3%), ST635 (6.4%) and ST10 (6.4%) predominated. Three isolates showed new allele combinations and three new STs (ST8252, ST8254 and ST8340) were assigned. The isolates belonged to six phylogenetic groups with a high occurrence of group A (32/78; 41%), D (17/78; 21.8%), B1 (16/78; 20.5%) and B2 (11/78; 14.1%), whereas only one isolate each was assigned to phylogenetic group C (1/78;

1.3%) and F (1/78; 1.3%). One isolate (strain R27c) was controversial as it was classified into group A but its ClermonTyper mash group was B1. The distribution of STs and phylogenetic groups are marked in the phylogenetic tree (Fig. 3) and Table S2.

3.6. WGS of CTX-resistant E. coli *– ARG, virulence genes, and plasmids*

Among the 78 sequenced *E. coli* isolates, ARG conferring resistance to beta-lactams (100%, n=78), macrolides (96.2%), aminoglycosides (74.4%), sulphonamides (70.5%), trimethoprim (61.3%), fluoroquinolones (50%), tetracyclines (39.8%) and phenicols (6.4%) were detected and are listed in Table S2. Beta-lactam resistance genes encoding ESBLs were predominant and mainly consisted of various $bla_{\text{CTX-M}}$ variants including $bla_{\text{CTX-M-15}}$ (62.8%), $bla_{\text{CTX-M-14}}$ (10.3%), $bla_{\text{CTX-M-1}}$ (9%), $bla_{\text{CTX-M-27}}$ (3.8%), $bla_{\text{CTX-M-3}}$ (2.6%), $bla_{\text{CTX-M-32}}$ (1.3%) and $bla_{\text{CTX-M-174}}$ (1.3%). Beta-lactam resistance genes encoding AmpC beta-lactamases were $bla_{\text{CMY-2}}$ (62.5%) and $bla_{\text{DHA-1}}$ (25%). Macrolide resistance genes identified were *mph*(A) in nineteen (24.4%) isolates and *erm*(B) in two (2.6%) isolates. Furthermore, *mdf*(A), which is considered to be intrinsic in *E. coli* (Edgar and Bibi, 1997), was detected in 75 (96.2%) isolates. At least one point mutation known to mediate quinolone resistances in chromosomal *gyrA* was found in 47.4% isolates (37/78) and at least one additional mutation in the genes *parC/parE* was detected in twenty-seven of these (34.6%). All isolates with at least three mutations (23/78) showed phenotypic resistance to nalidixic acid and ciprofloxacin (Table S2).

In regards to virulence-related genes, twenty-two isolates (28.2%) were predicted to be ExPEC and nearly half of the isolates originated from Municipal-WWTP-Out (10/22, 45.5%) (Table S2).

Thirty-one plasmid replicons were detected with FIB (48.7%), FII (43.6%), I1 (21.8%), FIB(K) (18%), FIA (15.4%), Y (15.4%), HI2 (14.1%), HI2A (12.8%) and Col156 (12.8%) being the most prevalent (Table S2).

3.7. Phylogenetic relatedness of selected E. coli

The phylogenetic tree (Fig. 3) was inferred based on nucleotide sequence aligment of 158,416 sites [i.e. positions (SNPs) that were detected in at least two isolates]. In general, the isolates with the same ST grouped together. The number of SNPs was in the range of 7 to 53,106 which confirmed high diversity relatedness between sequenced *E. coli* isolates (n=78). The isolates from the same phylogenetic groups were grouped together with the exception of strain R27c (River isolate 27c) which belonged to group A (ClermonType mash group was B1) and was the closest strain related to *E. coli* OM15c (Municipal-WWTP-Out isolate 15c) from group B1 (12,421 SNPs). The isolates with 100% identical PFGE profiles showed 324 SNPs [SH20c (Hospital-Raw isolate 20c) and IM9c (Municipal-WWTP-In isolate 15c)] and 799 SNPs [IM7c (Municipal-WWTP-In isolate 7c) and OM2c (Municipal-WWTP-Out isolate 2c)]. Few isolates showed relatedness ≤ 100 SNPs and they belonged to ST131 (4 strains from 2 different PFGE clusters), ST635 (3/3), ST38 (2/2), ST710 (2/1) and ST1970 (2/1). Individual phylogenetic trees for the most frequently detected STs, ST38 (n=10; 70,012 sites; SNPs in the range 75 – 12,112), ST131 (n=8; 88,075 sites; 71 – 13,259 SNPs), ST10 (n=5; 15.340 sites; 1,380 – 11,054 SNPs) and ST635 (n=5; 46,223 sites; 197 – 4,659 SNPs) were constructed (Fig. S1-S4).

3.8. Metagenomics analysis of water samples

The obtained reads for the metagenomic analysis of six water samples were in the range of 1.9 - 2.8 million. The percentage of mapped and unmapped reads (the reads which were not assigned to any of the databases) varied widely according to the origin of water sample with the highest number of mapped reads in Hospital-raw (82.9%) and the lowest in Municipal-

WWTP-Out (5.5%). Overall, most mapped reads (range: 5.3 - 81.7%) were assigned to bacteria (including waterborne *E. coli*–db, bacteria, bacteria draft and human microbiome databases) and low number of reads (reads average summed for each database < 0.2%) were assigned to protozoa, viruses, invertebrates, vertebrates-mammals, vertebrates-others and archaea. No reads were mapped to the human database. In the case of databases run in fullmode, a high abundance of mapped reads was observed only for the plasmid database (range: 0.2 - 13.8%). Low number of reads were mapped to ARG (0.08% on average), plasmid replicons (0.07% on average) and virulence genes (0.003% on average). Detailed information on number of reads mapped to the databases is listed in Table S3.

3.9. Bacterial composition of water samples

The bacterial composition varied between water samples. Overall, the majority of bacterial reads were assigned to the phylum Proteobacteria (range: 73 - 97%), followed by Bacteroidetes (2 - 15%) and Firmicutes (1 - 4%). The phyla Nitrospirae (9%), Actinobacteria (2%) and Chloroflexi (2%) were characteristic for Municipal-WWTP-Out and a high occurrence of the phylum Actinobacteria was observed in river water (19%) (Table S4a, S4b, S4c).

The predominant genera in Hospital-WWTP-In included *Aeromonas* (25.3%), *Acidovorax* (16%), *Escherichia* (14.2%), *Pseudomonas* (8.3%), *Acinetobacter* (6.3%), *Arcobacter* (4.1%) and *Klebsiella* (3.9%). In contrast, in Hospital-WWTP-Out *Acidovorax* (26.4%), *Aeromonas* (6.8%), *Thauera* (5.9%) were predominant and the remaining 60.9% of the bacterial composition was made by genera with an abundance of < 4%. The majority of genera that have pathogenic representatives were found in Hospital-Raw, including *Pseudomonas* (58.8%), *Escherichia* (9.8%), *Enterobacter* (8.6%), *Klebsiella* (7.5%) and *Aeromonas* (3.6%).

Municipal-WWTP-In included predominantly the genera *Arcobacter* (29.4%), *Aeromonas* (11.6%), *Acinetobacter* (9.3%), *Bacteroides* (8.5%), *Acidovorax* (7.3%) and *Citrobacter* (4.6%). The high occurrence (62.8%) of genera with relative abundance < 4% was observed in Municipal-WWTP-Out as well as the presence of genera such as *Acidovorax* (11.8%), *Nitrospira* (9.2%), *Arcobacter* (5.9%), *Thauera* (5.7%) and *Aeromonas* (4.6%).

The river sample exhibited a unique composition of major genera compared to the other samples with high relative abundance of *Limnohabitans* (41.1%), *Rhodoluna* (6.7%), *Candidatus* Planktophila (6.2%), *Polynucleobacter* (5.6%) and *Flavobacterium* (4.6%) (Fig. 4, Table S4a, S4b, S4c).

3.10. ARG and plasmids replicons in water samples

A total of 167 ARG clusters conferring resistance to compounds in 11 antimicrobial groups were detected across all water samples (Fig. 5a, 5b, Table S5a, S5b, S5c). The highest ARG abundance was documented in Hospital-WWTP-In (3,755 FPKM), followed by Hospital-Raw (595 FPKM), Municipal-WWTP-In (372 FPKM), and Hospital-WWTP-Out (336 FPKM). Low ARG abundances were observed in the Municipal-WWTP-Out and river (27 and 5 FPKM, respectively). High abundance of genes mediating resistance to beta-lactams (mainly encoded by *bla*$_{\text{GES}}$_clust, *bla*$_{\text{OXA}}$_clust2, *bla*$_{\text{OXA}}$_clust4), aminoglycosides (*aac(6')-Ib*_clust, *aac(6')-30-aac(6')-Ib'*, *aadA*_clust1, *aadA10*, *aph(6)-Id*) and aminoglycosides/quinolones (*aac(6')-Ib*_clust) were detected in hospital waters (Hospital-WWTP-In, Hospital-WWTP-Out, Hospital-Raw). Differently, resistance to macrolides, encoded mainly by genes *msr*(E) and *mph*(E), was most abundant in Municipal-WWTP-In. Notably, the gene *bla*$_{\text{CTX-M}}$ was not found in the metagenomic dataset.

The highest abundance of plasmid replicons was detected in Hospital-WWTP-In (12,167 FPKM), followed by Hospital-Raw (1,457 FPKM), Municipal-WWTP-In (790 FPKM),

Hospital-WWTP-Out (439 FPKM), Municipal-WWTP-Out (56 FPKM) and river (18 FPKM). The highest proportion of the 71 plasmid replicons belonged to small plasmids such as Col440I (detected in range 33.4 – 56.4% across the all water samples), ColKP3 (e.g. 16.1% in Hospital-WWTP-Out), ColRNAI (14.9% in Municipal-WWTP-Out) and IncQ2 (11.2% and 9.3% in Municipal-WWTP-In and Hospital-WWTP-In, respectively) and Col440II (5.1% in Hospital-Raw) (Fig. 5c, Table S6).

4. Discussion

Municipal and hospital WWs represent a rich source of different bacterial species in which antimicrobial resistant strains with pathogenic potential are frequently detected. Despite the fact that hospital WWs contain bacteria originating from the human population that is under medical treatment, only minimal contribution on content and diversity of ARG in influent to WWTP was observed (Buelow et al., 2018). This is probably caused by small amount (< 1%) of hospital WWs in municipal WWs (Karkman et al., 2018). For a more holistic understanding of the influence of hospital and community settings on spreading of AMR, analysis of both municipal and hospital WWs was performed in this study.

Cephalosporins (3rd, 4th and 5th generation), including cefotaxime, belong to the antimicrobials with highest priority on the WHO list of Critically Important Antimicrobials for Human Medicine (World Health Organisation, 2019) and thus the occurrence and spread of bacterial strains resistant to these antimicrobials represent a serious clinical threat. In this study, a high occurrence of CTX-resistant *E. coli* producing clinically important ESBLs was detected in hospital and municipal WWs including Municipal-WWTP-Out. Furthermore, high occurrence of several potentially pathogenic bacterial genera was observed in Hospital-WWTP-In, Hospital-Raw and Municipal-WWTP-In using a metagenomic approach.

The highest number of CTX-resistant *E. coli* and other coliforms were detected at Hospital-WWTP-In (wastewater from infection clinics, undetected after treatment), followed by Hospital-Raw and Municipal-WWTP-In. The abundance of bacteria monitored by cultivation decreased at Municipal-WWTP-Out compared to Municipal-WWTP-In, but the proportion of CTX-resistant *E. coli* and other coliform bacteria among their total counts was significantly higher in Municipal-WWTP-Out, likely suggesting higher tolerance of CTX-resistant bacteria to the wastewater treatment. Similar findings were documented previously in a study from France that reported a significantly higher proportion of ESBL-producing *E. coli* in treated (0.6%) compared to untreated (0.3%) water (Bréchet et al., 2014). These findings suggest an improved ability of ESBL-producing bacteria to survive treatment processes used in conventional WWTPs.

The phenotype and genotype analyses revealed a high occurrence of *E. coli* producing ESBLs (94.3%) encoded mostly by $bla_{\text{CTX-M}}$ (98.7% of 151 isolates) in our CTX-resistant *E. coli* collection (n=158). A low occurrence of CTX-resistant *E. coli* was observed in the river with only few ESBL- (n=4) and AmpC- (n=3) producing *E. coli*, compared to the other samples. Similarly, a high presence of $bla_{\text{CTX-M}}$ (96%) among 54 ESBL-producing *E. coli* was detected in the effluent from a municipal WWTP and in hospital WW in Japan (Gomi et al., 2017a). *E. coli* isolates carrying $bla_{\text{CTX-M}}$ were also documented in surface water in the Netherlands (Franz et al., 2015) and in the Yamato river in Japan (Gomi et al., 2017b). Another study conducted by Liu et al. (2018) characterized a total of 2,686 *E. coli* isolates from rivers and lakes in Northwest China and 2.8% were ESBL producers and 76.3% of 76 isolates carried a $bla_{\text{CTX-M}}$ gene. Despite the low occurrence of ESBL-producing *E. coli* in surface water, these observations are worrisome. They suggest the ability of these clinically important strains to become part of the natural environment, but conclusive studies connecting strains from WWTP effluent and nature are required.

Among the sequenced *E. coli* (n=78), a high genomic diversity (40 different STs) as well as large variety of ARG and plasmids were observed, yet some clinically important STs were dominating (ST38, ST131, ST635 and ST10). All *E. coli* ST131 (n=8) showed ExPEC status and carried $bla_{\text{CTX-M}}$ (mainly $bla_{\text{CTX-M-15}}$) which is in agreement with the pandemic spread of *E. coli* ST131 lineage and has been documented in water environments previously (Dolejska et al., 2011; Paulshus et al., 2019). A total of twenty-two *E. coli* isolates (28% of 78 isolates) had ExPEC status and interestingly ten of them originated from the outflow of the municipal WWTP. A similar finding was documented for a municipal WWTP in Germany where virulence profiling showed that 16% (of 50 isolates) and 19% (of 42 isolates) *E. coli* from inflow and outflow had ExPEC characteristics, respectively (Mahfouz et al., 2018).

The metagenomic analysis revealed differences in the bacterial composition among the water samples. The DNA isolation methods significantly influence observed microbial diversity (Knudsen et al., 2016; Zielińska et al., 2017) and thus the comparison with other studies with different methodology has to be made in caution. The three major phyla Proteobacteria (73 - 97%), Bacteroidetes (1 - 15%) and Firmicutes (1 - 4%) were predominantly found in our water samples as was documented in municipal waters previously (Narciso-da-Rocha et al., 2018). Narciso-da-Rocha et al. (2018) observed lower occurrence of Proteobacteria (66% and 43% in municipal WW and treated water, respectively), than what we noticed. In a study focused on surface water in China (Jin et al., 2018) the most predominant phyla were Proteobacteria (22 - 79%), Bacteroidetes (19 - 75%) and Actinobacteria (percentage was not specified). High occurrence of Actinobacteria (18%) was also detected in the river in our study and it seems to be typical for surface water (Jin et al., 2018).

The high relative abundance of potentially pathogenic bacteria (genera *Pseudomonas, Escherichia, Klebsiella, Enterobacter, Arcobacter* and *Aeromonas*) was characteristic for hospital WWs (Hospital-WWTP-In, Hospital-Raw) and Municipal-WWTP-In, compared to

Hospital-WWTP-Out and Municipal-WWTP-Out. In these samples, the majority of bacterial reads (over 84%) was represented by genera with relative abundance of < 4% and environmental bacteria (genera *Acidovorax*, *Thauera*, *Aeromonas* and *Nitrospira*). The high abundance of *Nitrospira* spp. (9%) observed at the outflow of the municipal WWTP is likely due to their ability to utilise nitrites in activated sludge (Gruber-Dorninger et al., 2015). This finding points out large changes in the bacterial composition during the wastewater treatment, both in the hospital and municipal WWTPs, thus corroborating previous findings (Bengtsson-Palme et al. 2016). Interestingly, the presence of *Escherichia* and other coliforms was observed in metagenomic results in all samples, but these bacteria had not been obtained by the cultivation approach from Hospital-WWTP-Out where chlorine disinfection is used. It was observed that only small portion of *E. coli* (0.4%) can survive the chlorine disinfection used in WWTP in the viable but nonculturable state (Oliver et al., 2005). This suggests the presence of DNA from coliform bacteria in Hospital-WWTP-Out, however whether this originated from live or dead bacteria is unknown.

The abundance of ARG was higher in the hospital and municipal WWs than in the treated waters and the river according to the metagenomic analysis. Similarly, a high proportion of genes mediating resistance to macrolides, beta-lactams, aminoglycosides and tetracyclines was detected in municipal sewage in Europe by Hendriksen et al. (2019). Surprisingly, the gene $bla_{\text{CTX-M}}$ was missing in our metagenomic dataset even though it was present in most of the isolates from our CTX-resistant *E. coli* collection. The possible explanation is probably insufficient sequencing depth of the water samples.

5. Conclusions

The obtained results revealed a widespread occurrence and high genetic diversity of ESBL-producing *E. coli* in hospital and municipal waters in the city of Brno, Czech Republic, based

on cultivation approach. ESBL-producing *E. coli* with ExPEC status were documented for several water samples, including the outflow from municipal WWTP. This water is flowing to the river and is used for irrigation and recreational activities. The subsequent contamination of surface water by clinically important bacteria represents a threat for human and animal health. The metagenomic approach revealed that the abundance and composition of bacteria, ARG and plasmids varied across tested water samples. This approach allowed to obtain a high-level perspective on the water samples including taxonomic shifts towards lower pathogenic potential of bacterial genera in WWTP outflows compared to inflows. The two approaches used in this study conveyed different information and highlighted the need for cautious interpretation of results. The culture-based approach was more appropriate to follow the dynamics of ESBL-producing *E. coli* compared to the metagenomics, which likely requires an increased sequencing depth to detect full composition of ARG in the water sample and to allow precision epidemiology. Taken together, the results of the two approaches convey information useful for public health risk assessment and suggest the need to improve wastewater treatment technology as a step to reduce spread of AMR to the environment.

This study brought above mentioned findings, however it would be beneficial to repeat sampling to observe changes in dynamic of ESBL-producing *E. coli*, bacterial composition and abundance of ARG over longer time period.

Acknowledgements

We would like to thank Jarmila Lausova, Eva Suchanova, Jana Hofirkova, Barbora Rettegy, Marco Minoia, Adam Valcek, Eva Tomcova, Lenka Dostalova, Marie Komarkova, Kristina Nesporova, Hassan Tarabai from University of Veterinary and Pharmaceutical Sciences Brno, Czech Republic, and Jacob D. Jensen and Kate V. Vibefeldt from the National Food Institute, Denmark for their excellent help in laboratory or data analysis. Our thanks go to Frank M.

Aarestrup (Head of Research Group for Genomic Epidemiology, National Food Institute, Denmark) who allowed the authors to perform the metagenomic analyses at the group. Furthermore, we thank the Core Facility Genomics of CEITEC for their support with whole-genome sequencing on NextSeq platform.

Data availability

Sequencing data are available from EMBL-EBI/NCBI/DDBJ under BioProject accession number PRJNA595483, the genome assemblies under accession numbers SAMN13556329 - SAMN13556406 (Table S2) and metagenomic data under accession numbers SAMN13558873 - SAMN13558878 (Table S3).

Appendix A. Supporting information

Supplementary data: Table S1-S6 and Fig. S1-S4.

References

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. J. Antimicrob. Chemother. 69, 1785–1791. https://doi.org/10.1093/jac/dku079

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455–477. https://doi.org/10.1089/cmb.2012.0021

Beghain, J., Bridier-Nahmias, A., Nagard, H. Le, Denamur, E., Clermont, O., 2018. ClermonTyping: an easy-to-use and accurate *in silico* method for *Escherichia* genus strain phylotyping. Microb. Genomics. 4, 1–8. https://doi.org/10.1099/mgen.0.000192

Bengtsson-Palme, J., Hammarén, R., Pal, C., Östman, M., Björlenius, B., Flach, C.-F., Fick, J., Kristiansson, E., Tysklind, M., Larsson, D.G.J., 2016. Elucidating selection processes for antibiotic resistance in sewage treatment plants using metagenomics. Sci. Total Environ. 572, 697–712. https://doi.org/10.1016/j.scitotenv.2016.06.228

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bréchet, C., Plantin, J., Sauget, M., Thouverez, M., Talon, D., Cholley, P., Guyeux, C., Hocquet, D., Bertrand, X., 2014. Wastewater treatment plants release large amounts of extended-spectrum β-lactamase-producing *Escherichia coli* into the environment. Clin.

Infect. Dis. 58, 1658–1665. https://doi.org/10.1093/cid/ciu190
Buelow, E., Bayjanov, J.R., Majoor, E., Willems, R.J.L., Bonten, M.J.M., Schmitt, H., van Schaik, W., 2018. Limited influence of hospital wastewater on the microbiome and resistome of wastewater in a community sewerage system. FEMS Microbiol. Ecol. 94, fiy087. https://doi.org/10.1093/femsec/fiy087
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 Microbiol. Ecol. 94, 1–14. https://doi.org/10.1093/femsec/fiy101
Carattoli, A., Zankari, E., García-Fernández, A., Voldby Larsen, M., Lund, O., Villa, L., Møller Aarestrup, F., Hasman, H., 2014. *In silico* detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob. Agents Chemother. 58, 3895–3903. https://doi.org/10.1128/aac.02412-14
Carrico, J.A., Pinto, F.R., Simas, C., Nunes, S., Sousa, N.G., Frazão, N., Lencastre, H. De, Almeida, J.S., 2005. Assessment of band-based similarity coefficients for automatic type and subtype classification of microbial isolates analyzed by pulsed-field gel electrophoresis. J. Clin. Microbiol. 43, 5483–5490. https://doi.org/10.1128/JCM.43.11.5483–5490.2005
Centers for Disease Control and Prevention. 2004. One-day (24–28 h) standardized laboratory protocol for molecular subtyping of *Escherichia coli* O157:H7, non-typhoidal *Salmonella* serotypes, and *Shigella sonnei* by pulsed field gel electrophoresis (PFGE). Training manual. US Department of Health and Human Services, Atlanta, GA.
CLSI, 2017. Performance standards for antimicrobial susceptibility testing, 27th ed. CLSI supplement M100. Wayne, PA: Clinical and Laboratory Standards Institute.
CLSI, 2015a. Performance standards for antimicrobial disk susceptibility tests, approved standard-twelfth edition. CLSI document M02-A12. Wayne, PA: Clinical and Laboratory Standards Institute.
CLSI, 2015b. Performance standards for antimicrobial susceptibility testing; twenty-fifth informational supplement. CLSI document M100-S25. Wayne, PA: Clinical and Laboratory Standards Institute.
CLSI, 2008. Performance standards for antimicrobial disk and dilution susceptibility tests for bacteria isolated from animals; approved standard-third edition. CLSI Document M31-A3. Wayne, PA: Clinical and Laboratory Standards Institute.
Correia, S., Poeta, P., Hébraud, M., Capelo, J.L., Igrejas, G., 2017. Mechanisms of quinolone action and resistance: where do we stand? J. Med. Microbiol. 66, 551–559. https://doi.org/10.1099/jmm.0.000475
Dale, A.P., Woodford, N., 2015. Extra-intestinal pathogenic *Escherichia coli* (ExPEC): Disease, carriage and clones. J. Infect. 71, 615–626. https://doi.org/10.1016/j.jinf.2015.09.009
Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2015. jModelTest 2: more models, new heuristics and high-performance computing. Nat. Methods 9, 772. https://doi.org/10.1038/nmeth.2109.
Day, A., 2015. heatmap.plus: Heatmap with more sensible behavior. R package version 1.3. https://CRAN.R-project.org/package=heatmap.plus 1–4.
Dobiasova, H., Dolejska, M., Jamborova, I., Brhelova, E., Blazkova, L., Papousek, I., Kozlova, M., Klimes, J., Cizek, A., Literak, I., 2013. Extended spectrum beta-lactamase and fluoroquinolone resistance genes and plasmids among *Escherichia coli* isolates from zoo animals, Czech Republic. FEMS Microbiol. Ecol. 85, 604–611. https://doi.org/10.1111/1574-6941.12149
Dolejska, M., Frolkova, P., Florek, M., Jamborova, I., Purgertova, M., Kutilova, I., Cizek, A.,

Guenther, S., Literak, I., 2011. CTX-M-15-producing *Escherichia coli* clone B2-O25b-ST131 and *Klebsiella* spp. isolates in municipal wastewater treatment plant effluents. J. Antimicrob. Chemother. 66, 2784–2790. https://doi.org/10.1093/jac/dkr363

Edgar, R., Bibi, E., 1997. MdfA, an *Escherichia coli* multidrug resistance protein with an extraordinarily broad spectrum of drug recognition. J. Bacteriol. 179, 2274–2280.

Franz, E., Veenman, C., Van Hoek, A.H.A.M., Husman, A.D.R., Blaak, H., 2015. Pathogenic *Escherichia coli* producing Extended-Spectrum β-Lactamases isolated from surface water and wastewater. Sci. Rep. 5, 1–9. https://doi.org/10.1038/srep14372

Garner, E., Chen, C., Xia, K., Bowers, J., Engelthaler, D.M., McLain, J., Edwards, M.A., Pruden, A., 2018. Metagenomic characterization of antibiotic resistance genes in full-scale reclaimed water distribution systems and corresponding potable systems. Environ. Sci. Technol. 52, 6113–6125. https://doi.org/10.1021/acs.est.7b05419

Garnier, S., 2018. viridis: default color maps from 'matplotlib'. R package version 0.5.1. https://cran.r-project.org/package=viridis.

Gomi, R., Matsuda, T., Matsumura, Y., Yamamoto, M., Tanaka, M., Ichiyama, S., Yonedaa, M., 2017a. Occurrence of clinically important lineages, including the sequence type 131 C1-M27 subclone, among extended-spectrum-β-lactamase-producing *Escherichia coli* in wastewater. Antimicrob. Agents Chemother. 61, e00564-17. https://doi.org/10.1128/AAC.00564-17

Gomi, R., Matsuda, T., Matsumura, Y., Yamamoto, M., Tanaka, M., Ichiyama, S., Yoneda, M., 2017b. Whole-genome analysis of antimicrobial- resistant and extraintestinal pathogenic *Escherichia coli* in river water. Appl. Environ. Microbiol. 83, e02703-16. https://doi.org/10.1128/AEM.02703-16

Gruber-Dorninger, C., Pester, M., Kitzinger, K., Savio, D.F., Loy, A., Rattei, T., Wagner, M., Daims, H., 2015. Functionally relevant diversity of closely related *Nitrospira* in activated sludge. ISME J. 9, 643–655. https://doi.org/10.1038/ismej.2014.156

Gweon, H.S., Shaw, L.P., Swann, J., De Maio, N., AbuOun, M., Niehus, R., Hubbard, A.T.M., Bowes, M.J., Bailey, M.J., Peto, T.E.A., Hoosdally, S.J., Walker, A.S., Sebra, R.P., Crook, D.W., Anjum, M.F., Read, D.S., Stoesser, N., The REHAB consortium, 2019. The impact of sequencing depth on the inferred taxonomic composition and AMR gene content of metagenomic samples. Environ. Microbiomes. 14, 1–15. https://doi.org/10.1186/s40793-019-0347-1

Hendriksen, R.S., Munk, P., Njage, P., van Bunnik, B., McNally, L., Lukjancenko, O., Röder, T., Nieuwenhuijse, D., Pedersen, S.K., Kjeldgaard, J., Kaas, R.S., Clausen, P.T.L.C., Vogt, J.K., Leekitcharoenphon, P., Schans, M.G.M. van de, Zuidema, T., Husman, A.M. de R., Rasmussen, S., Petersen, B., Consortium, the G.S.S. project, Amid, C., Cochrane, G., Sicheritz-Ponten, T., Schmitt, H., Alvarez, J.R.M., Aidara-Kane, A., Pamp, S.J., Lund, O., Hald, T., Woolhouse, M., Koopmans, M.P., Vigre, H., Petersen, T.N., Aarestrup, F.M., 2019. Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. Nat. Commun. 10, 1–12. https://doi.org/10.1038/s41467-019-08853-3

Jin, D., Kong, X., Cui, B., Jin, S., Xie, Y., Wang, X., Deng, Y., 2018. Bacterial communities and potential waterborne pathogens within the typical urban surface waters. Sci. Rep. 8, 1–9. https://doi.org/10.1038/s41598-018-31706-w

Joensen, K.G., Scheutz, F., Lund, O., Hasman, H., Kaas, R.S., Nielsen, E.M., Aarestrup, F.M., 2014. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. J. Clin. Microbiol. 52, 1501–1510. https://doi.org/10.1128/JCM.03617-13

Johnson, J.R., Menard, M., Johnston, B., Kuskowski, M.A., Nichol, K., Zhanel, G.G., 2009. Epidemic clonal groups of *Escherichia coli* as a cause of antimicrobial-resistant urinary

tract infections in Canada, 2002 to 2004. Antimicrob. Agents Chemother. 53, 2733–2739. https://doi.org/10.1128/AAC.00297-09

Karkman, A., Do, T.T., Walsh, F., Virta, M.P.J., 2018. Antibiotic-resistance genes in waste water. Trends Microbiol. 26, 220–228. https://doi.org/10.1016/j.tim.2017.09.005

Knudsen, B.E., Bergmark, L., Munk, P., Lukjancenko, O., Priemé, A., Aarestrup, F.M., Pamp, S.J., 2016. Impact of sample type and DNA isolation procedure on genomic inference of microbiome composition. mSystems 1, e00095-16. https://doi.org/10.1128/msystems.00095-16

Koboldt, D.C., Zhang, Q., Larson, D.E., Shen, D., McLellan, M.D., Lin, L., Miller, C.A., Mardis, E.R., Ding, L., Wilson, R.K., 2012. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res. 22, 568–576. https://doi.org/10.1101/gr.129684.111

Kolde, R., 2018. Pheatmap: pretty heatmaps. R package version 1.0.10. https://cran.r-project.org/package=pheatmap.

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359. https://doi.org/10.1038/nmeth.1923

Larsen, M. V., Cosentino, S., Rasmussen, S., Friis, C., Hasman, H., Marvig, R.L., Jelsbak, L., Sicheritz-Pontén, T., Ussery, D.W., Aarestrup, F.M., Lund, O., 2012. Multilocus sequence typing of total-genome-sequenced bacteria. J. Clin. Microbiol. 50, 1355–1361. https://doi.org/10.1128/JCM.06094-11

Letunic, I., Bork, P., 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res. 47, W256–W259. https://doi.org/10.1093/nar/gkz239

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup, 2009. The sequence alignment/map format and samtools. Bioinformatics 25, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Liu, B., Zheng, D., Jin, Q., Chen, L., Yang, J., 2019. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. Nucleic Acids Res. 47, D687–D692. https://doi.org/10.1093/nar/gky1080

Liu, H., Zhou, H., Li, Q., Peng, Q., Zhao, Q., Wang, J., Liu, X., 2018. Molecular characteristics of extended-spectrum β-lactamase-producing *Escherichia coli* isolated from the rivers and lakes in Northwest China. BMC Microbiol. 18, 1–12. https://doi.org/10.1186/s12866-018-1270-0

Mahfouz, N., Caucci, S., Achatz, E., Semmler, T., Guenther, S., Berendonk, T.U., Schroeder, M., 2018. High genomic diversity of multi-drug resistant wastewater *Escherichia coli*. Sci. Rep. 8, 1–12. https://doi.org/10.1038/s41598-018-27292-6

Manaia, C.M., Rocha, J., Scaccia, N., Marano, R., Radu, E., Biancullo, F., Cerqueira, F., Fortunato, G., Iakovides, I.C., Zammit, I., Kampouris, I., Vaz-Moreira, I., Nunes, O.C., 2018. Antibiotic resistance in wastewater treatment plants: Tackling the black box. Environ. Int. 115, 312–324. https://doi.org/10.1016/j.envint.2018.03.044

Narciso-da-Rocha, C., Rocha, J., Vaz-Moreira, I., Lira, F., Tamames, J., Henriques, I., Martinez, J.L., Manaia, C.M., 2018. Bacterial lineages putatively associated with the dissemination of antibiotic resistance genes in a full-scale urban wastewater treatment plant. Environ. Int. 118, 179–188. https://doi.org/10.1016/j.envint.2018.05.040

Oliver, J.D., Dagher, M., Linden, K., 2005. Induction of *Escherichia coli* and *Salmonella typhimurium* into the viable but nonculturable state following chlorination of wastewater. J. Water Health 3, 249–257. https://doi.org/10.2166/wh.2005.040

Pagès, H., Aboyoun, P., Gentleman, R., Debroy, S., 2019. Biostrings: efficient manipulation of biological strings. R package version 2.50.2. https://bioconductor.org/packages/release/bioc/html/Biostrings.html.

Paulshus, E., Thorell, K., Guzman-Otazo, J., Joffre, E., Colque, P., Kühn, I., Möllby, R., Sørum, H., Sjöling, Å., 2019. Repeated isolation of extended-spectrum-β-lactamase-positive *Escherichia coli* sequence types 648 and 131 from community wastewater indicates that sewage systems are important sources of emerging clones of antibiotic-resistant bacteria. Antimicrob. Agents Chemother. 63, e00823-19. https://doi.org/10.1128/AAC.00823-19

Peirano, G., Mulvey, G.L., Armstrong, G.D., Pitout, J.D.D., 2013. Virulence potential and adherence properties of *Escherichia coli* that produce CTX-M and NDM β-lactamases. J. Med. Microbiol. 62, 525–530. https://doi.org/10.1099/jmm.0.048983-0

Pérez-pérez, F.J., Hanson, N.D., 2002. Detection of plasmid-mediated AmpC β -lactamase genes in clinical isolates by using multiplex PCR. J. Clin. Microbiol. 40, 2153–2162. https://doi.org/10.1128/JCM.40.6.2153

Petersen, T.N., Lukjancenko, O., Thomsen, M.C.F., Sperotto, M.M., Lund, O., Aarestrup, F.M., Sicheritz-Pontén, T., 2017. MGmapper: reference based mapping and taxonomy annotation of metagenomics sequence reads. PLoS One. 12, 1–13. https://doi.org/10.1371/journal.pone.0176469

Pitout, J.D.D., Devinney, R., 2017. *Escherichia coli* ST131: a multidrug-resistant clone primed for global domination. F1000Research. 6, F1000 Faculty Rev-195. https://doi.org/10.12688/f1000research.10609.1

R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Russo, T.A., Johnson, J.R., 2003. Medical and economic impact of extraintestinal infections due to *Escherichia coli*: focus on an increasingly important endemic problem. Microbes Infect. 5, 449–456. https://doi.org/10.1016/S1286-4579(03)00049-2

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Warnes, G.R., Bolker, B., Bonebakker, L., Gentleman, R., Liaw, W.H.A., Lumley, T., Maechler, M., Magnusson, A., Moeller, S., Schwartz, M., Venables, B., 2019. gplots: various R programming tools for plotting data. R package version 3.0.1.1. http://CRAN.R-project.org/package=gplots.

World Health Organisation, 2019. WHO list of critically important antimicrobials for human medicine (WHO CIA list). World Heal. Organ. 6th revision.

Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., Aarestrup, F.M., Larsen, M. V., 2012. Identification of acquired antimicrobial resistance genes. J. Antimicrob. Chemother. 67, 2640–2644. https://doi.org/10.1093/jac/dks261

Zhou, Z., Alikhan, N.-F., Mohamed, K., Group, the A.S., Achtman, M., 2019. The user's guide to comparative genomics with EnteroBase. Three case studies: micro-clades within *Salmonella enterica* serovar Agama, ancient and modern populations of *Yersinia pestis*, and core genomic diversity of all *Escherichia*. bioRxiv 613554. https://doi.org/10.1101/613554

Zielińska, S., Radkowski, P., Blendowska, A., Ludwig-Gałęzowska, A., Łoś, J.M., Łoś, M., 2017. The choice of the DNA extraction method may influence the outcome of the soil microbial community structure analysis. Microbiologyopen. 6, 1–11. https://doi.org/10.1002/mbo3.453

Fig. 1 Water sampling sites

Six sites in Brno, Czech Republic were sampled within this study and are marked by red checkmarks: inflow to the hospital WWTP, outflow from the hospital WWTP, raw hospital sewage, inflow to the municipal WWTP, outflow from the municipal WWTP and the Svratka river.

Fig. 2 AMR phenotype, origin and detected beta-lactamases in CTX-resistant *E. coli* isolates (n=158)

The clustering of isolates is based on their AMR phenotype. AMR phenotype: AMP, ampicillin; KF, cephalotin; S, streptomycin; CAZ, ceftazidime; NA, nalidixic acid; SXT, sulphamethoxazole-trimethoprim; TE, tetracycline; AMC, amoxicillin-clavulanic acid; CN, gentamicin; S3, compound sulphonamides; CIP, ciprofloxacin; C, chloramphenicol; ETP, ertapenem; IMP, imipenem; MEM, meropenem. Strain origin: Inflow to the hospital WWTP (Hospital-WWTP-In), raw hospital sewage (Hospital-Raw), inflow to the municipal WWTP (Municipal-WWTP-In) and outflow from the municipal WWTP (Municipal-WWTP-Out).

Fig. 3 Phylogenetic tree of CTX-resistant *E. coli* isolates (n=78) based on SNPs analysis

See legend for the strain origin, phylogenetic group and variant of CTX-M beta-lactamase. Strain origin: Inflow to the hospital WWTP (Hospital-WWTP-In), raw hospital sewage (Hospital-Raw), inflow to the municipal WWTP (Municipal-WWTP-In) and outflow from the municipal WWTP (Municipal-WWTP-Out). The potential ExPEC are marked by black star. The strain identification (strain ID) and ST are denoted along the phylogenetic tree**.**

Fig. 4 Taxonomic composition of water samples at genus level

The relative abundance is presented as the percentage of each genus to the total read count for each water sample, respectively. The genera with relative abundance ≥ 4% at least in one of

the water samples are visualized, the remaining genera are included in the category "Other genera". [a]Inflow to the hospital WWTP (Hospital-WWTP-In), outflow from the hospital WWTP (Hospital-WWTP-Out), raw hospital sewage (Hospital-Raw), inflow to the municipal WWTP (Municipal-WWTP-In) and outflow from the municipal WWTP (Municipal-WWTP-Out). [b]Sum of percentage of reads mapped to genus *Escherichia* and database called waterborne *E. coli*. [c]Genus was not specified for this reference sequence in database, only phylum Actinobacteria was listed.

Fig. 5 Relative abundance (FPKM[c]) of ARG and plasmid replicons in water samples

a ARG abundance per antimicrobial class (ARG encoding resistance to the same antimicrobial group are clustered together). **b** Abundance of 30 most common ARG (ARG genes are clustered together based on the 95% similarity of reference sequences, the list of ARG for each gene cluster is listed in Table S6c). **c** Abundance of 30 most common plasmid replicons.

[a]Inflow to the hospital WWTP (Hospital-WWTP-In), outflow from the hospital WWTP (Hospital-WWTP-Out), raw hospital sewage (Hospital-Raw), inflow to the municipal WWTP (Municipal-WWTP-In) and outflow from the municipal WWTP (Municipal-WWTP-Out). [b]Only gene *aac(6')-Ib-cr* confers resistance to quinolones in the aac(6')-Ib_clust. [c]Colour strip shows fragments per kilo base per million fragments (FPKM) transformed to log (ln).

Highlights

- Antimicrobial resistance in wastewaters by culture-based and metagenomic approaches
- High prevalence of *E. coli* with $bla_{\text{CTX-M}}$ in municipal and hospital wastewaters
- ESBL-producing extraintestinal-pathogenic *E. coli* detected at outflow of WWTP
- Predominance of pathogenic bacteria in municipal and hospital wastewaters

Declaration of interests

☒ 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.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: