Whole Genome Sequencing of Extended-Spectrum Beta-Lactamase (ESBL)-Producing *Escherichia coli* Isolated From a Wastewater Treatment Plant in China

Xiawei Jiang[1], Xinjie Cui[1], Hao Xu[2], Wenhong Liu[1], Fangfang Tao[1], Tiejuan Shao[1], Xiaoping Pan[1] and Beiwen Zheng[2]*

[1] College of Basic Medical Sciences, Zhejiang Chinese Medical University, Hangzhou, China, [2] Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, The First Affiliated Hospital, College of Medicine, Zhejiang University, Hangzhou, China

Background and Objectives: Wastewater treatment plants (WWTPs) are one of the major reservoirs for antimicrobial resistant bacteria (ARB) and antimicrobial resistance genes (ARGs) in the environment. Thus, the investigation on ARB and ARGs from WWTPs has attracted increasing attention in recent years. In order to uncover the resistome in a WWTP treating effluents from a pharmaceutical industry in China, the extended-spectrum beta-lactamase (ESBL)-producing *Escherichia coli* strains were isolated and their whole genome sequences were obtained and analyzed. Moreover, metagenomic sequencing was applied to give a comprehensive view of antibiotic resistance in this WWTP.

Methods: 18 ESBL-producing *E. coli* strains were isolated from a WWTP located in Taizhou, China on April, 2017. All strains were sequenced using Illumina HiSeq 2000 sequencer. The whole genome sequences were assembled using SPAdes software and annotated with RAST server. Sequence types (STs), plasmids, ARGs and virulence genes were predicted from the genomes using MLST, Plasmid Finder, ResFinder and Virulence Finder, respectively. Metagenomic DNA of the same sample was extracted and sequenced using Illumina Hiseq X Ten platform. Metagenomic sequences were assembled using SOAPdenovo software.

Results: All 18 ESBL-producing *E. coli* strains were resistant to ampicillin, cefazolin, and ceftriaxone. Analysis of their genomes revealed that all strains carried beta-lactamase encoding genes and the most prevalent type was bla_{CTX-M}. Various virulence genes and ARGs confronting resistance to other types of antimicrobial agents were also predicted. Further investigation on the metagenomics data indicated 11 ARGs with high amino acid identities to the known ARGs. Five of these ARGs, *aadA1*, *aac(6′)-lb-cr*, *flo(R)*, *sul2* and *sul1*, were also present in the genomes of the ESBL-producing *E. coli* isolated from the same sample.

Conclusion: Our study revealed the resistome of a pharmaceutical WWTP by both culture-dependent and metegenomic methods. The existence of ESBL-producing *E. coli* strains, indicating that pharmaceutical WWTP can play a significant role in the emergence of ARB. The occurrence of ARGs annotated from the metagenomic data suggests that pharmaceutical WWTP can play a significant role in the emergence of ARGs. Our findings highlight the need for strengthening the active surveillance of ARB and ARGs from pharmaceutical industry.

Keywords: waste water treatment plants, ESBL-producing *Escherichia coli*, antimicrobial resistance genes, metagenomics, bla_{CTX-M}

INTRODUCTION

The rapid emergence and global spread of antimicrobial resistant bacteria (ARB) worldwide has long been recognized as a threat to human health, and one of the most studied bacterial groups is extended-spectrum beta-lactamase (ESBL)-producing Enterobacteriaceae (Bush and Fisher, 2011). An epidemiological study has reported that ESBL-producing Enterobacteriaceae, especially *Escherichia coli*, is an important cause of community-acquired infection (Stadler et al., 2018). ESBL enzymes produced by resistant strains are capable of inactivating beta-lactam drugs (e.g., third-generation cephalosporins) by hydrolyzing their beta-lactam rings (Bush and Bradford, 2016). Some of these ESBL-producing *E. coli* may also carry virulence genes. The products of these virulence genes enhance their pathogenicity. For instance, virulence genes *eilA*, *air*, and *lphA* encode for proteins that are related to adherence of bacteria (Doughty et al., 2002; Sheikh et al., 2006). Virulence genes *senB* and *astA* encode for enterotoxins (Dubreuil et al., 2016). Some virulence genes, such as *gad*, are involved in the survival of bacteria in harsh conditions such as the extremely acidic environments in stomach (Capitani et al., 2003). Thus, the infections caused by ESBL-producing *E. coli* strains which also harbor virulence genes are always associated with increased mortality and cost burden (Schwaber et al., 2006). In addition, ESBL genes are often encoded on plasmids that can be readily exchanged between species (Carattoli, 2013). Moreover, ESBL genes have been found not only in *E. coli* strains from clinical origin, but also from environmental origin (Runcharoen et al., 2017; Zheng et al., 2017a).

Wastewater treatment plants (WWTPs) have been regarded as vast reservoirs and environmental suppliers of ARB and antimicrobial resistance genes (ARGs) (Guo et al., 2017; Jałowiecki et al., 2018). Majority of the research has primarily focused on municipal WWTP receiving hospital and household wastewater (Yu et al., 2016; Szekeres et al., 2017; Galler et al., 2018). In recent years, the issue of pharmaceuticals in the environment has been gaining more attention worldwide and there has been growing interest in the WWTPs treating effluents from pharmaceutical industries with antimicrobial agents production. For example, several studies revealed the environmental impacts and the resistome of antimicrobial agents contaminated effluents from pharmaceutical industries in India and Croatia (Fick et al., 2009; Rutgersson et al., 2014;

Bielen et al., 2017; Razavi et al., 2017; González-Plaza et al., 2018). A recent study in Nigeria has applied a culture-depend method to isolate ARB in wastewater from pharmaceutical facilities (Obayiuwana et al., 2018). In China, studies have revealed the existence of ARGs in WWTPs that treated wastewater from antimicrobial agents producing factories (Li et al., 2010; Zhai et al., 2016; Tang et al., 2017). It is suggested that environmental pollution from pharmaceutical manufacturing can promote the spread of antimicrobial resistance, the induction of direct biostatic or bio-toxic effects to diverse organisms, and the alteration of species distribution (Bielen et al., 2017). Thus, there is a clear need for more research in this area.

The aim of the present study was to investigate the resistome in a WWTP of pharmaceutical industry in China. We used whole genome sequencing (WGS) to provide insight into genomic determinants of the ESBL-producing *E. coli* strains isolated from this WWTP and reveal their sequence types (STs), ARGs, virulence genes and plasmids. In addition, we applied a metagenomic approach to determine the microbial diversity and the occurrence of ARGs in this WWTP. The importance and originality of this study are derived from the identification of ARB and ARG in pharmaceutical WWTP using both WGS and metagenomic sequencing. Ultimately, it is hoped that this research will contribute to a deeper understanding of the ARB and ARGs released from pharmaceutical WWTP.

MATERIALS AND METHODS

Sample Collection

Sludge samples were collected from a WWTP located in Taizhou, China. This WWTP employed a conventional activated sludge process to treat wastewater from a beta-lactam antibiotics manufacturing company. A number of antimicrobial agents were produced by this company, including cefaclor, ceftizoxime, ceftibuten, cefuroxime sodium, cefprozil, cefdinir, and cefixime. There was an ongoing production at the time of sampling. Three independent samples (500 g for each sample) of aerobic active sludge were collected via the inspection opening of WWTP aeration tank in April 2017. All samples stored in sterilized tubes were transported to the laboratory on ice within 24 h.

Isolation and Antimicrobial Susceptibility Testing of ESBL-Producing *E. coli*

For pre-enrichment step, the samples were cultured overnight in Mueller-Hinton broth without antimicrobial agents. Subsequently, 500 μl enriched sample were spread on chromID® ESBL (bioMérieux, France) plates (the selective chromogenic medium for ESBL-producing enterobacteria screening). The plates were cultivated at 37°C under aerobic conditions. *E. coli* colonies were picked up from each plate according to the manufacturer's instructions (pink/burgundy) after 48 h. All these strains were re-cultivated on chromID® ESBL plates for confirmation. The identification of isolated bacteria was carried out using a Microflex MALDI-TOF mass spectrometer (Bruker Daltonics, Germany) as described previously (Zheng et al., 2017b). Antimicrobial susceptibility testing was determined on all strains by VITEK 2 system employing panel AST-GN-16 (bioMérieux, France). The results were then interpreted according to the standards of the Clinical and Laboratory Standards Institute (Clsi., 2017). *E. coli* ATCC 25922 was used as control.

Whole Genome Sequencing and Bioinformatic Analysis

Genomic DNA of the strains were extracted using QIAamp DNA Mini Kit (Qiagen, Germany) and were sent to Novogene Bioinformatics Technology, Co., Ltd. (Beijing, China) immediately on dry ice for WGS. Sequencing was carried out using the Illumina Hiseq 2000 sequencer (Illumina, United States) with a high-throughput 2 × 100 bp pair end sequencing strategy. Reads were filtered as described previously (Zheng et al., 2017b) and the resulting clean reads were assembled using SPAdes software (Bankevich et al., 2012). The draft genomes were annotated by RAST server (Aziz et al., 2008), while the 16S rRNA of strains were identified using EzBioCloud (Yoon et al., 2017). STs of the strains were predicted using MLST 2.0 (Larsen et al., 2012). Plasmids, ARGs and virulence genes were predicted from the genomes using PlasmidFinder 2.0 (Carattoli et al., 2014), ResFinder 3.1 (Zankari et al., 2012) and Virulence Finder 2.0 (Joensen et al., 2014) with nucleotide identity threshold of 90% and minimum length of 60%, respectively. Average nucleotide identity (ANI) between strains was calculated using their genomic sequences by Orthologous Average Nucleotide Identity Tool (OAT) with OrthoANI algorithm (Lee et al., 2016).

Metagenomic DNA Extraction

Metagenomic DNA was extracted from the samples upon arrival at the laboratory. DNA extraction was performed using FastDNA® Spin Kit for Soil (MP Biomedicals, Santa Ana, CA, United States) according to the manufacturer's instructions. The quality and integrity of DNA samples were checked by 1% agarose gel electrophoresis. Meanwhile, DNA yield and purity were measured using NanoDrop™ 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, United States) and Qubit® 2.0 fluorometer (Thermo Fisher Scientific, Waltham, MA, United States). After extraction, three DNA samples were mixed into one sample prior to metagenomic sequencing.

Metagenomic Sequencing and Bioinformatic Analysis

Metagenomic DNA was transported to Zhejiang Tianke High-Tech Development, Co., Ltd. (Hangzhou, China) on dry ice immediately after exaction. High-throughput sequencing was conducted using Illumina Hiseq Xten platform with PE150 (pair-end sequencing, 150 bp reads) sequencing mode. Raw reads generated from the sequencing were filtered by removing the reads with more than 40 bp low quality nucleotides, quality value lower than 38, more than 10 ambiguous nucleotides, and more than 15 bp overlapped with adapter. After filtering, the remaining clean reads were assembled using SOAPdenovo2 software (Luo et al., 2012). Different K-mer sizes (49, 55, and 59) were chosen for the assembly with the parameters of -d 1, -M 3, -u and -F, while the highest scaffold N50 value was selected for further analysis. The obtained scaffolds were then broken down at position N, resulting in the formation of continuous sequences within scaffolds called Scaftigs. Then, all the clean reads were aligned to Scaftigs using SOAPaligner software (Gu et al., 2013). Only Scaftigs longer than 500 bp were retained for further analysis. Open reading frames (ORFs) were predicted from the assembled Scaftigs using MetaGeneMark with default parameters (Zhu et al., 2010). CD-HIT was used to obtain a non-reduntant gene catalog (Fu et al., 2012). The retained clean reads were aligned to the gene catalog by using SOAPaligner, in order to calculate the number of matching reads (Gu et al., 2013). Unigenes were then obtained by removing the gene catalogs with less than three mapped reads. These unigenes were aligned against NCBI-NR database by DIAMOND software with the parameters of blastp, e-value ≤ 1e-5 (Buchfink et al., 2015) and further annotated with LCA algorithm using MEGAN (Huson et al., 2016). Further, ARGs were predicted by ResFinder 3.1 with 90% identity cutoff and 60% minimum length (Zankari et al., 2012).

RESULTS

Antimicrobial Resistance Profiles of *E. coli* Strains

A total of 18 pink colonies were picked up after spreading the samples on chromID® ESBL agar plates. By MALDI-TOF MS and 16S rRNA sequence analysis, all strains were identified as *E. coli*. As shown in **Table 1**, all strains were resistant to ampicillin, cefazolin and ceftriaxone, but not tigecycline.

Genomic Features of *E. coli* Strains

Draft genome sequences of 18 *E. coli* strains were obtained through WGS and assembling. The sequences had been deposited in DDBJ/ENA/GenBank under bioproject (PRJNA433283) with an individual accession number. Genome sizes ranged from 4.67 to 5.33 Mb, with an estimated GC content between 50.3 and 50.9%. A total of 4,576 to 5,240 coding sequences, 92 to 118 RNAs were annotated via RAST server. Sequence typing was

TABLE 1 | MIC values of all 18 *E. coli* strains.

Antibiotics (μg/ml)	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R16	R18	R19	R20
Ampicillin	≥32	≥32	≥32	≥32	≥32	≥32	≥32	≥32	≥32	≥32	≥32	≥32	≥32	≥32	≥32	≥32	≥32	≥32
Amoxicillin	8	4	≤2	16	16	4	16	16	4	4	≤2	8	16	8	4	4	8	16
Piperacillin	≤4	≤4	≤4	64	64	≤4	64	64	≤4	≤4	≤4	≤4	64	≤4	≤4	≤4	≤4	64
Cefazolin	≥64	≥64	≥64	≥64	≥64	≥64	≥64	≥64	≥64	≥64	≥64	≥64	≥64	≥64	≥64	≥64	≥64	≥64
Cefoxitin	≤4	≤4	≤4	64	64	16	64	64	≤4	8	≤4	≤4	64	≤4	16	16	≤4	64
Ceftriaxone	≥64	≥64	≥64	≥64	≥64	≥64	≥64	≥64	≥64	≥64	≥64	≥64	≥64	≥64	≥64	≥64	≥64	≥64
Cefepime	≤1	2	≤1	≥64	≥64	2	≥64	≥64	4	2	≤1	4	≥64	≤1	2	≥64	2	≥64
Aztreonam	≤1	2	2	≥64	≥64	2	≥64	≥64	4	16	≤1	4	≥64	16	16	16	4	≥64
Ertapenem	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5
Imipenem	≤1	≤1	≤1	≤1	≤1	≤1	≤1	≤1	≤1	≤1	≤1	≤1	≤1	≤1	≤1	≤1	≤1	≤1
Amikacin	8	≤2	≤4	≤2	≤2	≤2	≤2	≤2	≤2	≤2	≤2	≤2	≤2	≤2	≤2	≤2	≤2	≤2
Gentamicin	≥16	≤1	≤1	≥16	≤1	≥16	≥16	≤1	≤1	≤1	≤1	≤1	≤1	≥16	≤1	≥16	≤1	≤1
Tobramycin	≥16	≤1	≤1	8	≤1	8	8	≤1	≤1	≤1	≤1	≤1	≤1	≥16	≤1	≥16	≤1	≤1
Ciprofloxacin	≥4	≤0.25	≥4	≥4	≥4	0.5	≥4	≥4	≤0.25	≤0.25	≥4	≤0.25	≥4	1	≤0.25	≥4	≤0.25	≥4
Levofloxacin	≥8	≤0.25	4	4	≥8	1	≥8	≥8	≤0.25	≤0.25	≥8	≤0.25	≥8	1	≤0.25	≥8	≤0.25	≥8
Tigecycline	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5	≤0.5
Nitrofurantoin	≤16	≤16	64	≤16	≤16	≤16	≤16	≤16	≤16	≤16	≤16	≤16	≤16	≤16	≤16	128	32	≤16
Trimethoprim	≤20	≤20	≥320	≥320	≤20	≤20	≥320	≤20	≤20	≤20	≤20	≤20	≤20	≥320	≤20	80	≤20	≤20

The MIC values labeled with gray background indicate the resistance to the antimicrobial agents.

conducted with the whole genome sequences of the strains, generating eight different STs (**Table 2**). The strains that classified to ST405 and ST155 were accounted for 55.56% (*n* = 10). The STs represented by two strains were ST1193 (*n* = 2) and ST167 (*n* = 2). The remaining strains (*n* = 4) were singletons and represented their own ST.

ARGs in *E. coli* Strains

In order to detect the genes corresponding to the ESBL phenotype, we screened the genome sequences of the 18 *E. coli* strains for the presence of ARGs. The results indicated that all *E. coli* strains harbored bla_{CTX-M} gene (**Table 2**). Further analysis of the genetic environment of bla_{CTX-M} gene revealed that strains with same STs carried the same bla_{CTX-M} subtypes. For example, Strains R2, R9, R12 and R19, designated as ST155, harbored $bla_{CTX-M-14}$ gene, while strains R5, R10, R13 and R20, which were classified as ST405, carried two $bla_{CTX-M-15}$ genes and shared the same genetic environments (**Figure 1**). With such high similarities in gene arrangements among some strains, we wondered whether these strains were copy isolates identified in the samples. Therefore, we calculated the ANI values between the 18 isolates. The result showed that some strains shared 100% ANI, which might infer that they were actually copy isolates (**Supplementary Figure S1**). The genes encoding for mobile elements, transposases and plasmid conjugative transfer proteins were located near bla_{CTX-M} gene in majority of *E. coli* strains (**Figure 1** and **Supplementary Table S1**). Besides beta-lactamase genes, most strains also carried genes conferring resistance to many other classes of antimicrobial agents. For instance, 12 strains harbored aminoglycoside resistance genes with *aph(3″)-lb* and *aph(6)-ld* as the most abundant types; sulfonamide resistance genes (e.g., *sul1, sul2,* or *sul3*) were found in 12 strains. Eleven strains carried genes encoding for

tetracycline resistance, including *tet(A)*, *tet(B)*, *tet(D),* and *tet(M)*. Moreover, resistance genes conferring resistance to phenicols, trimethoprims, macrolides, lincosamides and streptogramin B (MLS), fluoroquinolones, and fosfomycins were detected in some of the strains.

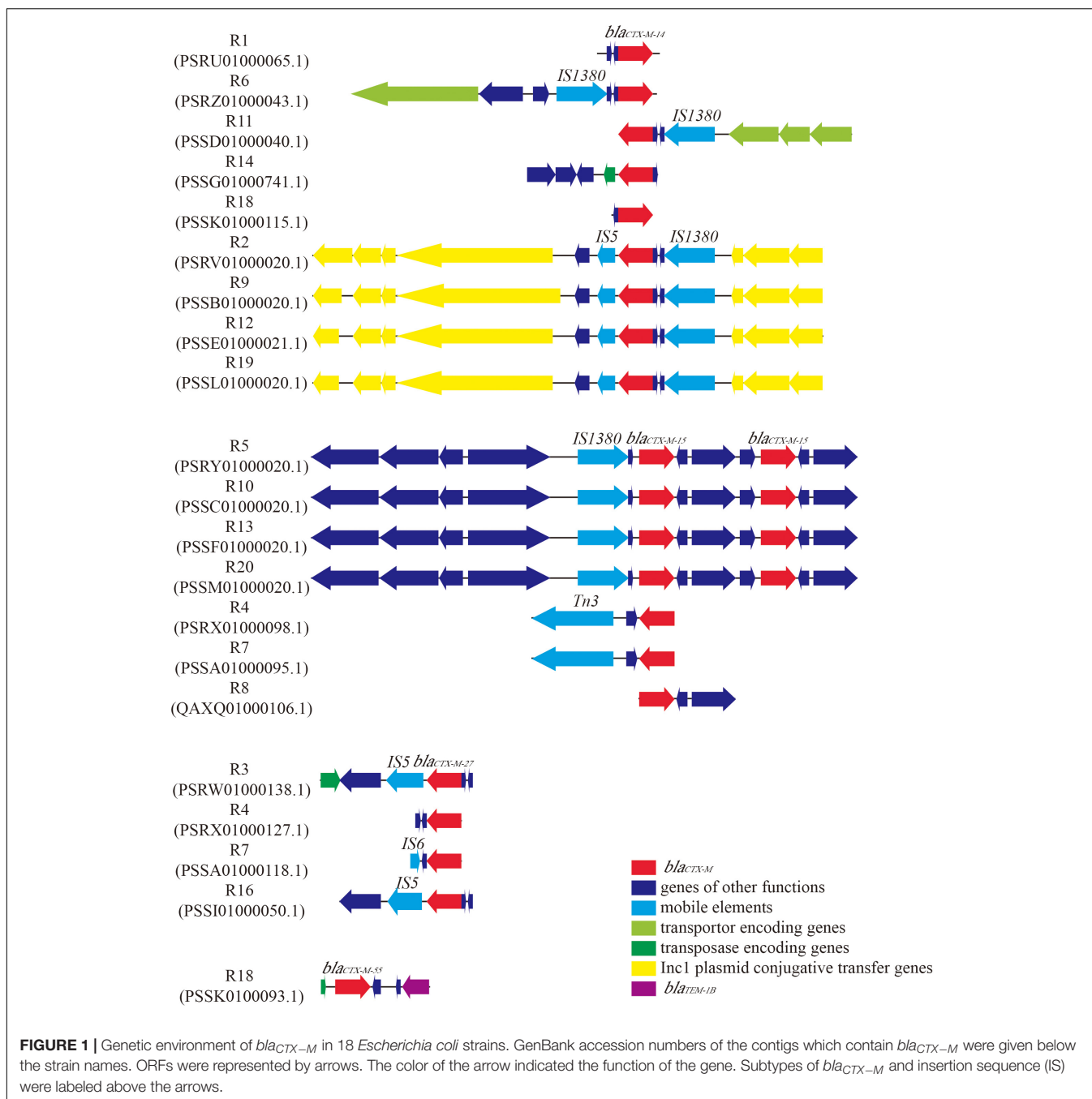## Virulence Genes in *E. coli* Strains

Virulence genes were detected in all *E. coli* strains and the numbers of virulence genes were different among the strains with different STs. One ST88 strain, R14, that encoded seven virulence genes (*gad*, *iroN*, *mchF*, *ireA*, *iss*, *celb*, and *tsh*), harbored more virulence genes than other STs strains. While the two ST167 strains, R4 and R7, only harbored *gad* gene. Of all the virulence genes, the glutamate decarboxylase encoding gene *gad* was the most common virulence gene presented in the 16 out of 18 strains (88.89%) (**Table 2**). The second predominant type was *eilA* (encoding for Salmonella HilA homolog), followed by *air* (encoding for Enteroaggregative immunoglobulin repeat protein) and *lpfA* (encoding for long polar fimbriae). Two enterotoxin genes, namely, *astA* and *senB*, were found in some *E. coli* strains.

Plasmids in *E. coli* Strains

17 out of 18 strains were predicted to harbor plasmids by PlasmidFinder (**Supplementary Table S2**). The Col156 plasmid replicon type was observed to be the most prevalent type among these strains (*n* = 13). The five *E. coli* ST405 strains (R5, R8, R10, R13, and R20) carry one Col(MP18) plasmid replicon type and one IncY plasmid incompatibility groups, with strain R5 and R13 additionally harboring one Col(156) replicon type. The four *E. coli* ST155 strains (R2, R9, R12, and R19) were identified to carry the same four plasmids: Col(MP18), Col(156), IncB/O/K/Z and IncY. However, another ST155 strain, R18, was identified

TABLE 2 | ARGs and virulence genes identified in the genomes of *E. coli* strains.

Strain	STs	Virulence genes	Acquired antimicrobial resistance genes of different antimicrobial classes								
			Beta-lactam	**Aminoglycoside**	**Fluoroquinolone**	**Fosfomycin**	**Macrolide, lincosamide and streptogramin B (MLS)**	**Phenicol**	**Sulfonamide**	**Tetracycline**	**Trimethoprim**
R1	ST1193	*gad senB vat iha sat*	$bla_{CTX-M-14}$ bla_{TEM-1B}	*aph(3″)-Ib aac(3)-IId aph(6)-Id aacA4*	*aac(6′)-Ib-cr*			*cmlA1*	*sul2*	*tet(A)*	
R2	ST155	*gad lpfA*	$bla_{CTX-M-14}$								
R3	ST450	*gad iha senB*	$bla_{CTX-M-27}$	*aph(3″)-Ib aadA5 aph(6)-Id*			*mph(A)*		*sul1 sul2*	*tet(A)*	*dfrA17*
R4	ST167	*gad*	$bla_{CTX-M-15}$ $bla_{CTX-M-27}$	*aadA5 aac(3)-IIa*			*mph(A) erm(B)*		*sul1*	*tet(B)*	*dfrA17*
R5	ST405	*gad eilA air*	$bla_{CTX-M-15}$ $bla_{CTX-M-15}$	*aph(3″)-Ib aph(6)-Id*				*floR*	*sul2*	*tet(A)*	
R6	ST38	*gad astA iss eilA*	$bla_{CTX-M-14}$	*aadA5 aac(3)-IId*			*mph(A)*		*sul1*		*dfrA17*
R7	ST167	*gad*	$bla_{CTX-M-15}$ $bla_{CTX-M-27}$	*aadA5 aac(3)-IIa*			*mph(A) erm(B)*		*sul1*	*tet(B)*	*dfrA17*
R8	ST405	*eilA air*	$bla_{CTX-M-15}$	*aph(3″)-Ib aph(6)-Id*				*floR*	*sul2*	*tet(A)*	
R9	ST155	*gad lpfA*	$bla_{CTX-M-14}$								
R10	ST405	*gad eilA air*	$bla_{CTX-M-15}$ $bla_{CTX-M-15}$	*aph(3″)-Ib aph(6)-Id*				*floR*	*sul2*	*tet(A)*	
R11	ST1193	*gad vat sat iha*	$bla_{CTX-M-14}$								
R12	ST155	*gad lpfA*	$bla_{CTX-M-14}$								
R13	ST405	*gad eilA air*	$bla_{CTX-M-15}$ $bla_{CTX-M-15}$	*aph(3″)-Ib aph(6)-Id*				*floR*	*sul2*	*tet(A)*	
R14	ST88	*gad iroN mchF ireA iss celb tsh*	$bla_{CTX-M-14}$	*aph(4)-Ia aac(3)-IVa aph(3′)-Ia aadA2 aadA1*		*fosA3*		*floR cmlA1*	*sul2 sul3*	*tet(B)*	*dfrA12*
R16	ST1722	*eilA air astA lpfA iss*	$bla_{CTX-M-27}$								
R18	ST155	*gad lpfA astA*	$bla_{CTX-M-14}$ $bla_{CTX-M-55}$ bla_{TEM-1B}	*aph(3″)-Ib aph(6)-Id aph(3′)-IIa aph(3′)-Ia aadA1 aph(4)-Ia aac(3)-IVa, aadA2*	*oqxA oqxB*	*fosA3*	*mph (A) erm (42)*	*floR cmlA1*	*sul1 sul2 sul3*	*tet(A) tet(M)*	*dfrA17*
R19	ST155	*gad lpfA*	$bla_{CTX-M-14}$								
R20	ST405	*gad eilA air*	$bla_{CTX-M-15}$ $bla_{CTX-M-15}$	*aph(3″)-Ib aph(6)-Id*				*floR*	*sul2*	*tet(A)*	

FIGURE 1 | Genetic environment of *bla*$_{CTX-M}$ in 18 *Escherichia coli* strains. GenBank accession numbers of the contigs which contain *bla*$_{CTX-M}$ were given below the strain names. ORFs were represented by arrows. The color of the arrow indicated the function of the gene. Subtypes of *bla*$_{CTX-M}$ and insertion sequence (IS) were labeled above the arrows.

to carry eight plasmids [Col(MG828), IncFII(pHN78A), IncHI2, IncHI2A, IncI1, IncIQ1, IncX1, and P0111].

ARGs in WWTP

By metagenomics sequencing, a total of 109,523,076 clean reads were generated and were then deposited in NCBI Sequence Read Archive with accession number of SRP132643. 11 ARGs with high identities were annotated using the assembled sequence generated by metagenomic sequencing. These ARGs are involved in the resistance to various types of antimicrobial agents including aminoglycosides, fluoroquinolones, macrolides,

phenicols, and sulfonamides (**Supplementary Table S3**). Five ARGs identified in metagenomic data were also carried by the *E. coli* strains isolated from the same WWTP (**Table 2**).

DISCUSSION

In recent years, a growing number of studies have shown that WWTPs may function as major reservoirs of ARB and ARGs (Conte et al., 2017; Guo et al., 2017; Szekeres et al., 2017). Although most studies have reported ARB and ARGs

from municipal and hospital WWTP, some researchers focused on the pharmaceutical industry WWTPs for the wastewater discharged from pharmaceutical industry may contain high concentrations of hazardous substances, including antimicrobial agents. High abundance of resistant strains has been observed in pharmaceutical WWTPs. In particular, 63.6 ± 6.4% of the bacterial strains isolated from penicillin production wastewater (Li et al., 2009) and 86% of the bacterial strains from a WWTP of multiple bulk drug manufacturers (Marathe et al., 2013) were resistant to many antimicrobial agents. Notwithstanding this, most of the bacterial strains isolated from these WWTPs were non-pathogenic. For instance, only two strains were identified as *E. coli* among the 189 strains isolated from a WWTP treating oxytetracycline, avermectin, and ivermectin production wastewater (Li et al., 2010). Therefore, in our study, we used selective plates for the isolation of ESBL-producing *E. coli* from the pharmaceutical WWTP. In total, we recovered 18 ESBL-producing *E. coli* strains. The results of antimicrobial susceptibility testing demonstrated that all strains displayed resistance to beta-lactams such as ampicillin, cefazolin, and ceftriaxone (**Table 1**).

In order to investigate the ARGs carried by these strains, WGS technology was applied to retrieve the draft genome sequences of all the 18 strains. Not surprisingly, all of the strains were found to carry at least one beta-lactamase gene, and the most prevalent ARGs type was bla_{CTX-M} (**Table 2**). Previous studies have demonstrated that bla_{CTX-M} is the most common ARGs type in ESBL-producing *E. coli* strains of human (Campos et al., 2018) and animal origin (Wu et al., 2018). The molecular homology analysis of bla_{CTX-M} positive *E. coli* strains further suggested a transmission of pathogens across water, food animals and human (Hu et al., 2013). Our finding in accordance with other studies and revealed that bla_{CTX-M} was the most dominant ARGs type in WWTPs of industrial origin. The results of sequence alignments indicated bla_{CTX-M}-containing contigs share a high degree of similarity among the strains with the same STs (**Figure 1** and **Table 2**). Further investigation of genetic environment revealed that the genes encoding for mobile elements, transporters, transposase and Incl1 plasmid conjugative transfer proteins located near the bla_{CTX-M} in most of the strains, suggesting the transfer potential of bla_{CTX-M} (**Figure 1**). In addition, plasmid prediction revealed that $bla_{CTX-M-14}$ genes and IncB/O/K/Z plasmids were located in the same contigs of strains R2, R9, R12, and R19. Likewise, bla_{TEM-1B} gene and plasmid IncFIA were predicted in the same contig of strain R1. Further indicating the transfer potential of these ARGs via plasmids.

Numerous virulence genes were detected in the genomes of the 18 ESBL-producing *E. coli* strains. For instance, all ST405 strains (R5, R8, R10, R13, and R20) harbored *eilA* and *air* genes which encoded for Salmonella HilA homolog and Enteroaggregative immunoglobulin repeat protein, respectively. The long polar fimbriae encoding gene *lpfA* was detected in all ST155 strains (R2, R9, R12, R18, and R19). These genes were demonstrated to play a role in the adherence of the bacteria (Doughty et al., 2002; Sheikh et al., 2006). Strain R14 which belonged to ST88, harbored the highest number of virulence genes among all the strains. The possession of a high

number of virulence traits in ST88 stains was also reported in previous studies (Oteo et al., 2014). Among these virulence genes, *gad* was the most prevalent type (**Table 2**). The *gad* gene enables the bacteria to survive in harsh acidic conditions in gastric environment by helping them to maintain a near-neutral intracellular pH (Capitani et al., 2003). Genes encoded for enterotoxins were also discovered. Strain R1 and R3 carried *senB* gene which encoded for plasmid-encoded enterotoxin. While strain R6, R16, and R18 harbored the heat-stable enterotoxin 1 encoding gene *astA*. These enterotoxins are responsible for fluid secretion and may associated with diarrhea (Dubreuil et al., 2016). With all these virulence genes, the *E. coli* strains isolated from the pharmaceutical WWTP may pose a potential threat to human health.

In addition to metagenomic sequencing, numerous studies have employed quantitative polymerase chain reaction (qPCR) for the detection of ARGs and their relative abundance (Szekeres et al., 2017; Zheng et al., 2017c). However, due to the limited availability of qPCR primers, metagenomic sequencing may be more appropriate for the screening of both known and novel ARGs (Schmieder and Edwards, 2012; Berglund et al., 2017; Marathe et al., 2018). Nevertheless, a combination of WGS and metagenomic sequencing can be an effective strategy to unravel the resistome in WWTP. To further reveal ARGs in this WWTP, a high-throughput metagenomic sequencing technology was implemented. The occurrence of 11 ARGs was analyzed from the metegenomic sequence with high amino acid identities (**Supplementary Table S3**). Five of these ARGs, *aadA1*, *aac(6')-lb-cr*, *flo(R)*, *sul2* and *sul1*, were also present in the genomes of the *E. coli* isolated from the same WWTP. However, we failed to find any beta-lactam resistance genes in the metagenomic data, while these beta-lactam resistance genes were carried by all the 18 *E. coli* isolated from the same sample. That may due to the low abundance of these beta-lactam resistance genes in the metagenomics data. By taxonomy annotation, only three out of 105,892 unigenes were annotated as *E. coli*, indicating that *E. coli* was of very low abundance in the WWTP. This low abundance of *E. coli* in WWTP was also observed in previous study (Li et al., 2010). Only those *E. coli* that exhibit antimicrobial resistance can survive in this WWTP which treats the antimicrobial agents producing pharmaceutical wastewater. The present work reported the isolation and analysis of ESBL-producing *E. coli* in a pharmaceutical WWTP and presented their antimicrobial phenotypes as well as genotypes. However, the background of *E. coli* in this geographic area is still unknown. In order to draw a whole picture of the prevalence of ESBL-producing *E. coli* and to evaluate the potential risk of their transmission in this area, further study needs to be done to elucidate the diversity of *E. coli* in this area.

CONCLUSION

In present study, we applied both WGS and metagenomics sequencing technologies to uncover the antimicrobial resistance profiles of ESBL-producing *E. coli* strains and the antimicrobial resistome in a pharmaceutical WWTP. Our findings indicate

that pharmaceutical WWTP can play a significant role in the emergence of ARB and ARGs. In order to prevent the spread of ARB and ARGs, the sewage sludge produced by this WWTP are treated through dewatering and incineration. Thus, further research is required to investigate the ARB and ARGs in the sludge dewatering streams.

AUTHOR CONTRIBUTIONS

XJ and BZ designed the study. XC and HX performed the experiments. WL, FT, TS, and XP contributed reagents, materials, and analysis tools. XJ and BZ analyzed the data and wrote the manuscript. All authors revised the manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2019.01797/full#supplementary-material

REFERENCES

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., et al. (2008). The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9:75. doi: 10.1186/1471-2164-9-75

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021

Berglund, F., Marathe, N. P., Österlund, T., Bengtsson-Palme, J., Kotsakis, S., Flach, C.-F., et al. (2017). Identification of 76 novel B1 metallo-β-lactamases through large-scale screening of genomic and metagenomic data. *Microbiome* 5:134. doi: 10.1186/s40168-017-0353-8

Bielen, A., Šimatoviæ, A., Kosiæ-Vukšiæ, J., Senta, I., Ahel, M., Babiæ, S., et al. (2017). Negative environmental impacts of antibiotic-contaminated effluents from pharmaceutical industries. *Water Res.* 126, 79–87. doi: 10.1016/j.watres.2017.09.019

Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12:59. doi: 10.1038/nmeth.3176

Bush, K., and Bradford, P. A. (2016). β-Lactams and β-lactamase inhibitors: an overview. *Cold Spring Harbor Pers. Med.* 6:a025247. doi: 10.1101/cshperspect.a025247

Bush, K., and Fisher, J. F. (2011). Epidemiological expansion, structural studies, and clinical challenges of new β-lactamases from gram-negative bacteria. *Ann. Rev. Microbiol.* 65, 455–478. doi: 10.1146/annurev-micro-090110-102911

Campos, A. C. C., Andrade, N. L., Ferdous, M., Chlebowicz, M. A., Santos, C. C., Correal, J. C., et al. (2018). Comprehensive molecular characterization of *escherichia coli* isolates from urine samples of hospitalized patients in rio de janeiro. Brazil. *Front. Microbiol.* 9:243. doi: 10.3389/fmicb.2018.00243

Capitani, G., De Biase, D., Aurizi, C., Gut, H., Bossa, F., and Grütter, M. G. (2003). Crystal structure and functional analysis of *Escherichia coli* glutamate decarboxylase. *EMBO J.* 22, 4027–4037. doi: 10.1093/emboj/cdg403

Carattoli, A. (2013). Plasmids and the spread of resistance. *Int. J. Med. Microbiol.* 303, 298–304. doi: 10.1016/j.ijmm.2013.02.001

Carattoli, A., Zankari, E., García-Fernandez, A., Larsen, M. V., Lund, O., Villa, L., et al. (2014). PlasmidFinder and pMLST: in silico detection and typing of plasmids. *Antimicro. Agents Chemother.* 58, 3895–3903. doi: 10.1128/AAC.02412-14

Clsi. (2017). *Performance Standards for Antimicrobial Susceptibility Testing*, 27th Edn. Wayne, PA: CLSI.

Conte, D., Palmeiro, J. K., da Silva Nogueira, K., de Lima, T. M. R., Cardoso, M. A., Pontarolo, R., et al. (2017). Characterization of CTX-M enzymes, quinolone resistance determinants, and antimicrobial residues from hospital sewage, wastewater treatment plant, and river water. *Ecotoxicol. Environ. Safe.* 136, 62–69. doi: 10.1016/j.ecoenv.2016.10.031

Doughty, S., Sloan, J., Bennett-Wood, V., Robertson, M., Robins-Browne, R. M., and Hartland, E. L. (2002). Identification of a novel fimbrial gene cluster related to long polar fimbriae in locus of enterocyte effacement-negative strains of enterohemorrhagic *Escherichia coli*. *Infect. Immun.* 70, 6761–6769. doi: 10.1128/iai.70.12.6761-6769.2002

Dubreuil, J., Isaacson, R., and Schifferli, D. (2016). Animal Enterotoxigenic *Escherichia coli*. *EcoSal Plus* 7. doi: 10.1128/ecosalplus.ESP-0006-2016

Fick, J., Söderström, H., Lindberg, R. H., Phan, C., Tysklind, M., and Larsson, D. J. (2009). Contamination of surface, ground, and drinking water from pharmaceutical production. *Environ. Toxicol. Chem.* 28, 2522–2527. doi: 10.1897/09-073.1

Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28, 3150–3152. doi: 10.1093/bioinformatics/bts565

Galler, H., Feierl, G., Petternel, C., Reinthaler, F. F., Haas, D., Habib, J., et al. (2018). Multiresistant bacteria isolated from activated sludge in austria. *Int. J. Environ. Res. Public Health* 15:479. doi: 10.3390/ijerph15030479

González-Plaza, J. J., Šimatoviæ, A., Milakoviæ, M., Bielen, A., Wichmann, F., Udikoviæ-Koliæ, N., et al. (2018). Functional repertoire of antibiotic resistance genes in antibiotic manufacturing effluents and receiving freshwater sediments. *Front. Microbiol.* 8:2675. doi: 10.3389/fmicb.2017.02675

Gu, S., Fang, L., and Xu, X. (2013). Using SOAPaligner for short reads alignment. *Curr. Protoc. Bioinformat.* 44, 11.11.11–11.11.17. doi: 10.1002/0471250953.bi1111s44

Guo, J., Li, J., Chen, H., Bond, P. L., and Yuan, Z. (2017). Metagenomic analysis reveals wastewater treatment plants as hotspots of antibiotic resistance genes and mobile genetic elements. *Water Res.* 123, 468–478. doi: 10.1016/j.watres.2017.07.002

Hu, Y. Y., Cai, J. C., Zhou, H. W., Chi, D., Zhang, X. F., Chen, W. L., et al. (2013). Molecular typing of CTX-M-producing *Escherichia coli* isolates from environmental water, swine feces, specimens from healthy humans, and human patients. *Appl. Environ. Microbiol.* 79, 5988–5996. doi: 10.1128/AEM.01740-13

Huson, D. H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., et al. (2016). MEGAN community edition-interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS Comput. Biol.* 12:e1004957. doi: 10.1371/journal.pcbi.1004957

Jałowiecki, Ł., Żur, J., Chojniak, J., Ejhed, H., and Płaza, G. (2018). Properties of antibiotic-resistant bacteria isolated from onsite wastewater treatment plant in relation to biofilm formation. *Curr. Microbiol.* 75, 639–649. doi: 10.1007/s00284-017-1428-2

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

Larsen, M. V., Cosentino, S., Rasmussen, S., Friis, C., Hasman, H., Marvig, R. L., et al. (2012). Multilocus sequence typing of total-genome-sequenced bacteria. *J. Clin. Microbiol.* 50, 1355–1361. doi: 10.1128/JCM.06094-11

Lee, I., Kim, Y. O., Park, S. C., and Chun, J. (2016). OrthoANI: an improved algorithm and software for calculating average nucleotide identity. *International J. Syst. Evol. Microbiol.* 66, 1100–1103. doi: 10.1099/ijsem.0.000760

Li, D., Yang, M., Hu, J., Zhang, J., Liu, R., Gu, X., et al. (2009). Antibiotic-resistance profile in environmental bacteria isolated from penicillin production wastewater treatment plant and the receiving river. *Environ. Microbiol.* 11, 1506–1517. doi: 10.1111/j.1462-2920.2009.01878.x

Li, D., Yu, T., Zhang, Y., Yang, M., Li, Z., Liu, R., et al. (2010). Antibiotic resistance characteristics of environmental bacteria from an oxytetracycline production wastewater treatment plant and the receiving river. *Applied and Environ. Microbiol.* 76, 3444–3451. doi: 10.1128/AEM.02964-09

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., et al. (2012). SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience* 1:18. doi: 10.1186/2047-217X-1-18

Marathe, N. P., Janzon, A., Kotsakis, S. D., Flach, C.-F., Razavi, M., Berglund, F., et al. (2018). Functional metagenomics reveals a novel carbapenem-hydrolyzing mobile beta-lactamase from Indian river sediments contaminated with antibiotic production waste. *Environ. Int.* 112, 279–286. doi: 10.1016/j.envint.2017.12.036

Marathe, N. P., Regina, V. R., Walujkar, S. A., Charan, S. S., Moore, E. R., Larsson, D. J., et al. (2013). A treatment plant receiving waste water from multiple bulk drug manufacturers is a reservoir for highly multi-drug resistant integron-bearing bacteria. *PLoS One* 8:e77310. doi: 10.1371/journal.pone.0077310

Obayiuwana, A., Ogunjobi, A., Yang, M., and Ibekwe, M. (2018). Characterization of bacterial communities and their antibiotic resistance profiles in wastewaters obtained from pharmaceutical facilities in lagos and ogun states, nigeria. *Int. Environ. Res. Public Health* 15:1365. doi: 10.3390/ijerph15071365

Oteo, J., González-López, J. J., Ortega, A., Quintero-Zárate, J. N., Bou, G., Cercenado, E., et al. (2014). Inhibitor-resistant TEM-and OXA-1-producing *Escherichia coli* isolates resistant to amoxicillin-clavulanate are more clonal and possess lower virulence gene content than susceptible clinical isolates. *Antimicrob. Agents Chemother.* 58, 3874–3881. doi: 10.1128/AAC.02738-13

Razavi, M., Marathe, N. P., Gillings, M. R., Flach, C.-F., Kristiansson, E., and Larsson, D. J. (2017). Discovery of the fourth mobile sulfonamide resistance gene. *Microbiome* 5:160. doi: 10.1186/s40168-017-0379-y

Runcharoen, C., Raven, K. E., Reuter, S., Kallonen, T., Paksanont, S., Thammachote, J., et al. (2017). Whole genome sequencing of ESBL-producing *Escherichia coli* isolated from patients, farm waste and canals in Thailand. *Genome Med.* 9:81. doi: 10.1186/s13073-017-0471-8

Rutgersson, C., Fick, J., Marathe, N., Kristiansson, E., Janzon, A., Angelin, M., et al. (2014). Fluoroquinolones and qnr genes in sediment, water, soil, and human fecal flora in an environment polluted by manufacturing discharges. *Environ. Sci. Technol.* 48, 7825–7832. doi: 10.1021/es501452a

Schmieder, R., and Edwards, R. (2012). Insights into antibiotic resistance through metagenomic approaches. *Future Microbiol.* 7, 73–89. doi: 10.2217/fmb.11.135

Schwaber, M. J., Navon-Venezia, S., Kaye, K. S., Ben-Ami, R., Schwartz, D., and Carmeli, Y. (2006). Clinical and economic impact of bacteremia with extended-spectrum-β-lactamase-producing *Enterobacteriaceae*. *Antimicro. Agents Chemother.* 50, 1257–1262. doi: 10.1128/aac.50.4.1257-1262.2006

Sheikh, J., Dudley, E. G., Sui, B., Tamboura, B., Suleman, A., and Nataro, J. P. (2006). EilA, a HilA-like regulator in enteroaggregative *Escherichia coli*. *Mol. Microbiol.* 61, 338–350. doi: 10.1111/j.1365-2958.2006.05234.x

Stadler, T., Meinel, D., Aguilar-Bultet, L., Huisman, J. S., Schindler, R., Egli, A., et al. (2018). Transmission of ESBL-producing *Enterobacteriaceae* and their mobile genetic elements—identification of sources by whole genome sequencing: study protocol for an observational study in Switzerland. *BMJ Open* 8:e021823. doi: 10.1136/bmjopen-2018-021823

Szekeres, E., Baricz, A., Chiriac, C. M., Farkas, A., Opris, O., Soran, M.-L., et al. (2017). Abundance of antibiotics, antibiotic resistance genes and bacterial community composition in wastewater effluents from different romanian hospitals. *Environ. Poll.* 225, 304–315. doi: 10.1016/j.envpol.2017.01.054

Tang, M., Dou, X., Wang, C., Tian, Z., Yang, M., and Zhang, Y. (2017). Abundance and distribution of antibiotic resistance genes in a full-scale anaerobic–aerobic system alternately treating ribostamycin, spiramycin and paromomycin production wastewater. *Environ. Geochem. Health* 39, 1595–1605. doi: 10.1007/s10653-017-9987-5

Wu, C., Wang, Y., Shi, X., Wang, S., Ren, H., Shen, Z., et al. (2018). Rapid rise of the ESBL and mcr-1 genes in *Escherichia coli* of chicken origin in China, 2008–2014. *Emerg. Microbes Infect.* 7:30. doi: 10.1038/s41426-018-0033-1

Yoon, S. H., Ha, S. M., Kwon, S., Lim, J., Kim, Y., Seo, H., et al. (2017). Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int. J. Syst. Evol. Microbiol.* 67, 1613–1617. doi: 10.1099/ijsem.0.001755

Yu, H., Qu, F., Shan, B., Huang, B., Jia, W., Chen, C., et al. (2016). Detection of the mcr-1 colistin resistance gene in carbapenem-resistant *Enterobacteriaceae* from different hospitals in China. *Antimicro. Agents Chemother.* 60, 5033–5035. doi: 10.1128/AAC.00440-16

Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., et al. (2012). Identification of acquired antimicrobial resistance genes. *J. Antimicro. Chemother.* 67, 2640–2644. doi: 10.1093/jac/dks261

Zhai, W., Yang, F., Mao, D., and Luo, Y. (2016). Fate and removal of various antibiotic resistance genes in typical pharmaceutical wastewater treatment systems. *Environ. Sci. Poll. Res.* 23, 12030–12038. doi: 10.1007/s11356-016-6350-9

Zheng, B., Huang, C., Xu, H., Guo, L., Zhang, J., Wang, X., et al. (2017a). Occurrence and genomic characterization of ESBL-Producing, MCR-1-Harboring *Escherichia coli* in Farming Soil. *Fron. Microbiol.* 8:2510.

Zheng, B., Jiang, X., Cheng, H., Guo, L., Zhang, J., Xu, H., et al. (2017b). Genome characterization of two bile-isolated Vibrio fluvialis strains: an insight into pathogenicity and bile salt adaption. *Sci. Rep.* 7:11827. doi: 10.1038/s41598-017-12304-8

Zheng, J., Gao, R., Wei, Y., Chen, T., Fan, J., Zhou, Z., et al. (2017c). High-throughput profiling and analysis of antibiotic resistance genes in east tiaoxi river. *China Environ. Poll.* 230, 648–654. doi: 10.1016/j.envpol.2017.07.025

Zhu, W., Lomsadze, A., and Borodovsky, M. (2010). Ab initio gene identification in metagenomic sequences. *Nucleic Acids Res.* 38, e132–e132. doi: 10.1093/nar/gkq275