Whole-Genome Analysis of Antimicrobial-Resistant and Extraintestinal Pathogenic *Escherichia coli* in River Water

Ryota Gomi,^a Tomonari Matsuda,^b Yasufumi Matsumura,^c Masaki Yamamoto,^c Michio Tanaka,^c Satoshi Ichiyama,^c Minoru Yoneda^a

Department of Environmental Engineering, Graduate School of Engineering, Kyoto University, Kyoto, Japan^a; Research Center for Environmental Quality Management, Kyoto University, Otsu, Japan^b; Department of Clinical Laboratory Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan^c

ABSTRACT Contamination of surface waters by antimicrobial-resistant bacteria and pathogenic bacteria is a great concern. In this study, 531 *Escherichia coli* isolates obtained from the Yamato River in Japan were evaluated phenotypically for resistance to 25 antimicrobials. Seventy-six isolates (14.3%) were multidrug resistant (MDR), 66 (12.4%) were nonsusceptible to one or two classes of agents, and 389 (73.3%) were susceptible. We performed whole-genome sequencing of selected strains by using Illumina technology. In total, the genome sequences of 155 strains were analyzed for antibiotic resistance determinants and phylogenetic characteristics. More than 50 different resistance determinants, including acquired resistance genes and chromosomal resistance mutations, were detected. Among the sequenced MDR strains ($n = 66$), sequence type 155 (ST155) complex ($n = 9$), ST10 complex ($n = 9$), and ST69 complex ($n = 7$) were prevalent. Among extraintestinal pathogenic *E. coli* (ExPEC) strains ($n = 58$), clinically important clonal groups, namely, ST95 complex ($n = 18$), ST127 complex ($n = 8$), ST12 complex ($n = 6$), ST14 complex ($n = 6$), and ST131 complex ($n = 6$), were prevalent, demonstrating the clonal distribution of environmental ExPEC strains. Typing of the *fimH* (type 1 fimbrial adhesin) gene revealed that ST131 complex strains carried *fimH22* or *fimH41*, and no strains belonging to the *fimH30* subgroup were detected. Fine-scale phylogenetic analysis and virulence gene content analysis of strains belonging to the ST95 complex (one of the major clonal ExPEC groups causing community-onset infections) revealed no significant differences between environmental and clinical strains. The results indicate contamination of surface waters by *E. coli* strains belonging to clinically important clonal groups.

IMPORTANCE The prevalence of antimicrobial-resistant and pathogenic *E. coli* strains in surface waters is a concern because surface waters are used as sources for drinking water, irrigation, and recreational purposes. In this study, MDR and ExPEC strains in river water were characterized by genomic sequencing and analysis. We detected more than 50 resistance determinants and identified clonal groups specific to MDR and ExPEC strains. This study showed contamination of surface waters by *E. coli* strains belonging to clinically important clonal groups. Overall, this study advances our understanding of environmental MDR and ExPEC strains.

KEYWORDS antimicrobial resistance, *Escherichia coli*, ExPEC, whole-genome sequencing

Escherichia coli is a commensal member of the gastrointestinal tract of humans and other warm-blooded animals. Consequently, *E. coli*, as well as *Enterococcus* spp., has long been used as an indicator of fecal contamination to assess the microbial quality of surface waters. Although most *E. coli* strains are harmless, some strains are pathogenic

Received 28 September 2016 Accepted 9 December 2016

Accepted manuscript posted online 16 December 2016

Citation Gomi R, Matsuda T, Matsumura Y, Yamamoto M, Tanaka M, Ichiyama S, Yoneda M. 2017. 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>.

Editor Frank E. Löffler, University of Tennessee and Oak Ridge National Laboratory

Copyright © 2017 American Society for Microbiology. All Rights Reserved.

Address correspondence to Tomonari Matsuda, matsuda.tomonari.8z@kyoto-u.ac.jp.

and increase the risk of waterborne diseases (1). Pathogenic *E. coli* can be classified as intestinal pathogenic *E. coli* or extraintestinal pathogenic *E. coli* (ExPEC) (2). Recently, we determined that most potential pathogenic *E. coli* isolates obtained from the Yamato River in Japan are ExPEC strains, based on their virulence gene profiles (3). Indeed, environmental water is considered to be one of the potential sources of ExPEC strains (4). Clinical ExPEC strains have often been characterized at the sequence type (ST) level by multilocus sequence typing (MLST), and STs such as ST73, ST95, and ST131 were found to be predominant among such isolates (5). However, limited information exists regarding the fine-scale phylogeny and virulence gene profiles of ExPEC strains in surface waters.

Antimicrobial-resistant *E. coli* is another concern, because antimicrobial resistance is a key factor in emerging infectious diseases (6). *E. coli* strains can easily gain resistance to antimicrobials and serve as vectors for dissemination of antimicrobial resistance (7). One of the main mechanisms involved in the spread of antimicrobial resistance is the emergence and dissemination of specific clones that acquire resistance determinants (8). Clonal group A (9), which belongs to ST69 and frequently harbors trimethoprim-sulfamethoxazole resistance, and ST131 H30-Rx (10), which produces extended-spectrum β -lactamase (ESBL), are good examples of such clones. Multidrug-resistant (MDR) clones, which are usually defined as those resistant to three or more drug classes, particularly are a great concern because MDR *E. coli* tends to harbor various resistance genes and to transfer its resistance determinants to other strains, species, or genera (8). Moreover, infections caused by MDR bacteria can lead to inadequate or delayed antimicrobial therapy (11). Importantly, many studies have reported the presence of MDR *E. coli* strains in surface waters and water-related environments, on the basis of phenotypic resistance testing and detection of several resistance determinants (12–18). Some studies also investigated the clonal composition of MDR *E. coli* strains based on MLST (13, 14). However, most of those studies targeted MDR isolates carrying specific resistance genes (e.g., ESBL genes) (13), and few studies have investigated the clonal composition of any type of MDR isolates (14). Moreover, information on comprehensive resistance determinant profiles of drug-resistant *E. coli* in surface waters is lacking.

In a recent study of *E. coli* strains obtained from the Yamato River, we determined that 12% of the strains were ExPEC (3). Although we sequenced the whole genomes of some ExPEC strains, clonal composition was not investigated. Moreover, we did not evaluate antimicrobial susceptibility of the isolated *E. coli* strains. In the present study, we performed antimicrobial susceptibility tests with 531 *E. coli* strains to determine the prevalence of MDR strains. We then sequenced and analyzed the whole genomes of selected strains to investigate the clonal composition of MDR strains and ExPEC strains in surface waters. We also used the sequence data to identify comprehensive resistance determinant profiles of environmental *E. coli* isolates. A whole-genome approach was employed because it provides comprehensive information on genetic characteristics such as virulence genes, antibiotic resistance determinants, and phylogeny.

RESULTS AND DISCUSSION

Antimicrobial susceptibility. The bacteria were tested for susceptibility to 25 antimicrobials. Assessing the susceptibility of environmental bacteria according to clinical breakpoints is inadequate because clinical breakpoints are based on parameters relevant to therapeutic success. In contrast, EUCAST epidemiological cutoff (ECOFF) values are epidemiologically based and thus are more suitable for interpreting findings for environmental bacteria (19). Therefore, results in the present study were interpreted according to ECOFF values as well as Clinical and Laboratory Standards Institute (CLSI) criteria. The antimicrobial susceptibilities of 531 isolates are shown in Fig. 1. Nonsusceptibility rates for the antimicrobial agents ranged from 0% (imipenem, meropenem, and colistin) to 16.8% (ampicillin). Carbapenems and colistin are used as antibiotics of last resort (20), and isolates resistant to these antimicrobials have rarely been identified in surface waters. The distribution of resistance patterns is shown in Fig. 2. The rates of susceptible isolates were high in our study, compared with other studies that examined

FIG 1 Antimicrobial susceptibilities of 531 isolates. Isolates showing nonsusceptibility according to ECOFF criteria or resistance according to CLSI criteria are shown as nonsusceptible isolates. Nonsusceptibility is grouped by categories, separated by vertical lines, in the following order: penicillins, penicillins plus β -lactamase inhibitors, antipseudomonal penicillins plus β -lactamase inhibitors, broad-spectrum first-generation cephalosporins, extended-spectrum third- and fourth-generation cephalosporins, cephamycins, carbapenems, monobactams, fluoroquinolones, aminoglycosides, folate pathway inhibitors, tetracyclines, phenicols, polymyxins, phosphonic acids, and others. AMP, ampicillin; AMC, amoxicillin-clavulanic acid; SAM, ampicillin-sulbactam; TZP, piperacillin-tazobactam; CFZ, cefazolin; CTX, cefotaxime; CAZ, ceftazidime; FEP, cefepime; FOX, ceftiofur; IPM, imipenem; MEM, meropenem; ATM, aztreonam; CIP, ciprofloxacin; GEN, gentamicin; TOB, tobramycin; AMK, amikacin; KAN, kanamycin; SXT, trimethoprim-sulfamethoxazole; TET, tetracycline; MIN, minocycline; CHL, chloramphenicol; CST, colistin; FOF, fosfomycin; CPD, cefpodoxime; NAL, nalidixic acid.

E. coli isolates in surface waters of different countries (17, 21), even though we also interpreted the results based on ECOFF values, which are usually lower than CLSI breakpoints. One possible reason for this finding is the difference in antibiotic use in medicine and agriculture. We found that the rates of antimicrobial nonsusceptibility for ExPEC isolates were higher than those for non-ExPEC isolates. ExPEC strains are sometimes the predominant *E. coli* type in the feces of certain human populations (22, 23); this might have increased the probability of ExPEC strains encountering antibiotics and contributed to the relatively high rates of antimicrobial nonsusceptibility. This is just one possible explanation, however, and further studies are needed to determine the actual sources of ExPEC strains.

Resistance determinants. Overall, more than 50 different resistance determinants were detected in 155 isolates (Table 1). The prevalence of resistance determinants among the different groups is shown in Table S1 in the supplemental material. As expected, no resistance determinants were detected in susceptible isolates. All of the isolates with ESBL genes, acquired *ampC* genes, or chromosomal *ampC* promoter/

FIG 2 Distribution of resistance patterns. ExPEC strains include only nonredundant and sequenced strains ($n = 58$). MDR was defined as nonsusceptible to one or more agents in three or more classes.

TABLE 1 Resistance determinants detected in 155 isolates

Antibiotic class ^a	Resistance determinant (no. of isolates) ^b
Aminoglycoside	<i>strA</i> (35), <i>strB</i> (35), <i>aadA1</i> (7), <i>aadA2</i> (5), <i>aadA5</i> (14), <i>aac(3)-IIa</i> (1), <i>aac(3)-IIId</i> (10), <i>aac(3)-VIa</i> (2), <i>aac(6')-Ia</i> (1), <i>aac(6')-Ib</i> (1), <i>aac(6')-Ib-cr</i> (1), <i>aph(3')-Ia</i> (5), <i>aph(3')-Ic</i> (1), <i>aph(3')-IIa</i> (1)
β -Lactam	<i>bla</i> _{TEM-1} (49), <i>bla</i> _{CARB-2} (1), <i>bla</i> _{OXA-1} (1), <i>bla</i> _{CMY-2} (6), <i>bla</i> _{SHV-2a} (1), <i>bla</i> _{SHV-12} (3), <i>bla</i> _{CTX-M-1} (1), <i>bla</i> _{CTX-M-2} (2), <i>bla</i> _{CTX-M-3} (1), <i>bla</i> _{CTX-M-8} (1), <i>bla</i> _{CTX-M-14} (7), <i>bla</i> _{CTX-M-15} (2), <i>bla</i> _{CTX-M-24} (2), <i>bla</i> _{CTX-M-55} (2), <i>ampC</i> promoter/attenuator mutation(s) (5)
Quinolone	<i>qnrB7</i> (3), <i>qnrD</i> (5), <i>qnrS1</i> (3), <i>qepA</i> (1), <i>aac(6')-Ib-cr</i> (1), mutation(s) in QRDRs (39)
Fosfomycin	<i>fosA</i> (2)
Macrolide-lincosamide-streptogramin B	<i>ere(B)</i> (1), <i>erm(B)</i> (3), <i>mph(A)</i> (21), <i>mph(B)</i> (1)
Phenicol	<i>cmlA1</i> (3), <i>catA1</i> (7), <i>catB3</i> (1), <i>floR</i> (9)
Sulfonamide	<i>sul1</i> (24), <i>sul2</i> (36), <i>sul3</i> (3)
Tetracycline	<i>tet(A)</i> (33), <i>tet(B)</i> (17), <i>tet(M)</i> (1)
Trimethoprim	<i>dfrA1</i> (3), <i>dfrA5</i> (2), <i>dfrA7</i> (1), <i>dfrA8</i> (2), <i>dfrA12</i> (4), <i>dfrA14</i> (6), <i>dfrA17</i> (15), <i>dfrB4</i> (1)

^aDrugs corresponding to each antibiotic class used in our study are as follows: aminoglycosides, gentamicin, tobramycin, amikacin, and kanamycin; beta-lactams, ampicillin, cefazolin, cefotaxime, ceftazidime, cefepime, cefoxitin, imipenem, meropenem, aztreonam, amoxicillin-clavulanic acid, ampicillin-sulbactam, piperacillin-tazobactam, and cefpodoxime; quinolones, nalidixic acid and ciprofloxacin; fosfomycins, fosfomycin; phenicols, chloramphenicol; sulfonamides and trimethoprim, trimethoprim-sulfamethoxazole; tetracyclines, tetracycline and minocycline.

^bThe numbers of isolates carrying each resistance determinant are presented in parentheses.

attenuator mutations were classified as MDR because these isolates were nonsusceptible to at least three antimicrobial categories related to β -lactams.

Fourteen different aminoglycoside resistance genes were detected. The *strA* and *strB* genes, which confer streptomycin resistance, were the most prevalent and were always detected together. Other streptomycin resistance genes, i.e., *aadA1*, *aadA2*, and *aadA5*, were also detected at relatively high frequencies. We did not test susceptibility to streptomycin because this agent is not used to define MDR (11). The *aac(3)-IIId* gene, which confers resistance to gentamicin and tobramycin, was detected in 10 isolates and was the most prevalent aminoglycoside resistance gene other than the streptomycin resistance genes. The *aac(6')-Ib-cr* gene, an aminoglycoside resistance gene that also reduces the activity of ciprofloxacin (24), was detected in one isolate.

Fourteen types of β -lactam resistance genes were detected. The *bla*_{TEM-1} gene was found in 49 isolates and was the most common. CTX-M-type ESBLs were more prevalent than SHV-type ESBLs, and no other types of ESBLs were found in the sequenced isolates. The *bla*_{CTX-M-14} gene was the most prevalent ESBL gene. Interestingly, this gene was also reported to be a common ESBL gene in clinical *E. coli* isolates collected in adjacent regions in Japan (25). The *bla*_{CMY-2} gene, which is the most common plasmid-mediated *ampC* β -lactamase gene worldwide (26), was the only acquired *ampC* β -lactamase gene detected in this study. Mutations in the chromosomal *ampC* promoter/attenuator regions that can result in *ampC* overexpression were detected in five isolates. Three ST12 isolates carrying *fimH447* had the -32 mutation and two ST6088 isolates had the -42 mutation in the *ampC* promoter/attenuator regions.

Fluoroquinolone resistance is conferred mainly by mutations in the *gyrA* and *parC* genes (27). The numbers of mutations leading to amino acid substitutions in quinolone resistance-determining regions (QRDRs) were closely associated with ciprofloxacin and nalidixic acid resistance; isolates with no mutations were susceptible to both agents, isolates with one or two mutations were resistant to nalidixic acid but not to ciprofloxacin (nonsusceptible to ciprofloxacin according to the ECOFF value), and isolates with three or four mutations were resistant to both agents. These results are consistent with those of studies that analyzed clinical *E. coli* isolates (27, 28). Plasmid-mediated quinolone resistance (PMQR) determinants, i.e., *qnrB7*, *qnrD*, *qnrS1*, *qepA*, and *aac(6')-*

lb-cr, were detected in 10 isolates. Eight isolates with *qnr* genes did not carry mutations in QRDRs and were nonsusceptible to ciprofloxacin according to the ECOFF value but were susceptible to nalidixic acid. This finding is consistent with a previous report (29) and suggests that interpreting results with ECOFF values improves the detection of resistance mechanisms, because isolates carrying PMQR genes are still susceptible to ciprofloxacin according to the clinical breakpoint.

Fosfomycin resistance mechanisms include decreased drug uptake, modification of the target site, and inactivation of the antibiotic (30). The first and second mechanisms are explained by chromosomal mutations, while the third mechanism is explained by the acquisition of enzymes that inactivate the antibiotic. However, the relationships between these mechanisms and fosfomycin nonsusceptibility of *E. coli* based on ECOFF criteria have not been well investigated. Five of 531 isolates investigated in the present study were nonsusceptible to fosfomycin according to the ECOFF value, and one isolate was even resistant according to the clinical breakpoint. This relatively low prevalence of fosfomycin resistance may be attributable to the biological cost of the mutations that confer resistance to this agent (30). The single fosfomycin-resistant isolate (according to the clinical breakpoint) carried *fosA*. This gene was also detected in a susceptible isolate but was truncated. The distribution of fosfomycin-resistant *E. coli* in surface waters is not well known. One previous study reported fosfomycin-resistant *E. coli* isolates in urban waterways, but the resistance determinants were not investigated (18).

E. coli is intrinsically resistant to macrolide antibiotics (31). Interestingly, we detected macrolide resistance genes, i.e., *ere*(B), *erm*(B), *mph*(A), and *mph*(B), in 23 isolates. The *erm*(B), *mph*(A), and *mph*(B) genes were also detected in *E. coli* in a previous study (32). These results suggest that *E. coli* may be a major reservoir for macrolide resistance genes, which could be subsequently transferred to other bacteria.

We detected four different chloramphenicol resistance genes, i.e., *cmlA1*, *catA1*, *catB3*, and *floR*. The same four genes were detected in chloramphenicol-resistant *E. coli* strains in a previous study, which also reported *floR* as the most prevalent gene (33).

Resistance to sulfonamides in *E. coli* results mainly from the acquisition of an alternative dihydropteroate synthase gene (*sul*) (34). Three types of *sul* genes, i.e., *sul1*, *sul2*, and *sul3*, have been reported (35), and all three were detected in the present study; *sul1* and *sul2* were relatively prevalent, while *sul3* was detected in only three isolates. The frequencies of detection of these three *sul* genes seem to differ among studies (15, 17), demonstrating the diversity of *sul* gene distribution profiles in various areas.

Three types of tetracycline resistance genes, i.e., *tet*(A), *tet*(B), and *tet*(M), were detected in 48 isolates. Two isolates carried both *tet*(A) and *tet*(B), and one isolate carried both *tet*(A) and *tet*(M). Four of 30 *tet*(A)-carrying isolates were resistant to both tetracycline and minocycline, while the other 26 were resistant to tetracycline but susceptible to minocycline. In contrast, 16 of 17 isolates carrying *tet*(B) were resistant to both tetracycline and minocycline, which reflects the unique ability of Tet(B) to extrude both tetracycline and minocycline (36). A previous study similarly reported that *tet*(A) was more prevalent than *tet*(B) and *tet*(M) in *E. coli* strains isolated from surface waters (15).

Trimethoprim resistance is explained mainly by the acquisition of a trimethoprim-insensitive dihydrofolate reductase, which is the target enzyme of this agent (37). To date, more than 30 different *dfr* genes are known (35). Eight of these *dfr* genes were detected in the present study, and *dfrA17* was the most prevalent variant. This variant has been found frequently in clinical *E. coli* isolates and is known to be present as a gene cassette within class 1 integrons (37).

Overall, resistance genotypes correlated well with resistance phenotypes (see Data Set S1 in the supplemental material). We did not investigate resistance mechanisms such as porins or efflux pumps, however, because associations of these mechanisms with phenotypic resistance for some agents are not completely understood. These mechanisms may explain the phenotype-genotype discrepancies observed for some strains.

The study area is densely populated, and several wastewater treatment plants (WWTPs) and hospitals are present along the Yamato River. Animal husbandry facilities as well as some WWTPs and hospitals are located in the upstream part of the river system. These are the probable contamination sources of antimicrobial resistance genes and antimicrobial-resistant *E. coli* detected in this study, although further studies are needed to confirm this.

PlasmidFinder identified various types of plasmid replicons (Data Set S1). Although we were able to link resistance genes directly to replicon sequences in some cases (e.g., *qnrD* with Col3M), correlations between plasmid replicons and resistance genes could not be established in many cases, mainly because the current sequencing and *de novo* assembly methods are not able to assemble complete plasmid sequences (38); this does not necessarily mean that the resistance genes are unrelated to the plasmids carrying the detected replicons. We could not perform integron analysis using our draft genome sequences because complete integrons, *In0* elements, and clusters of *attC* sites lacking integron-integrases (CALIN) cannot be accurately identified when contig breaks occur in these sequences (39).

Diversity of genotypes. A whole-genome single nucleotide polymorphism (SNP)-based tree was constructed using kSNP3.0. Initially, we included assembled genomes from all isolates and reference genomes to assign each strain to a phylogenetic group or a cryptic clade. A table that was provided in a previous study (40) and described relationships between phylogenetic groups and STs was also considered for the assignment. We found that 150 isolates belonged to one of the six phylogenetic groups (group A, B1, B2, C, D, or F) and one isolate belonged to cryptic clade V. The phylogenetic group memberships of four *E. coli* strains, namely, KOr026, KKa024, KKo009, and KFu027, were ambiguous. Cryptic lineages are novel lineages of the genus *Escherichia* that are genetically distinct but phenotypically indistinguishable from *E. coli* (41). One previous study reported a low frequency of resistance for cryptic clades (42), and our isolate belonging to clade V was susceptible to all of the agents tested. Next, we constructed a phylogenetic tree after removing some reference strains and our strain belonging to clade V (Fig. 3). The topology of the tree in Fig. 3 was very similar to that of a tree constructed on the basis of all SNPs, which was the default tree generated by kSNP (data not shown). It is well understood that bacterial recombination occurs frequently and reduces phylogenetic signals (43). Moreover, the efficiency with which kSNP detects SNPs depends on the diversity within the data set (44). In order to assess the effects of recombination and diversity on the accuracy of the tree, the ratio of r/m (the amount of change introduced by recombination relative to the amount introduced by mutation) (45) to the fraction of core kmers was calculated; for this data set, the ratio was 3.31. It has been reported that the probability that a kSNP tree will be at least 90% accurate is >0.9 when this ratio is less than 5 (44). Therefore, the tree in Fig. 3 can be considered to be accurate and reliable.

Among 155 isolates typed by MLST, 80 STs were identified, including 11 novel STs. Table 2 shows the distribution of the major STs and phylogenetic groups of the 155 isolates. We found that 54 (93%) of 58 ExPEC isolates belonged to phylogenetic group B2 or D. Interestingly, all ExPEC isolates that were nonsusceptible to one or two classes of agents and all susceptible ExPEC isolates belonged to phylogenetic group B2. In contrast, only 22 (23%) of 97 non-ExPEC isolates belonged to these two phylogenetic groups. These results are consistent with the fact that phylogenetic groups B2 and D are typically associated with human ExPEC infections (22). MDR isolates were relatively dispersed in terms of phylogenetic groups. However, some clonal groups, namely, ST complex 69 (STC69) and STC155, were characterized by MDR isolates. These points are discussed further below.

Characteristics of MDR strains. As pointed out above, STC69 and STC155 were characterized by MDR isolates. In addition, STC10 contained nine MDR isolates. However, the isolates belonging to STC10 were somewhat heterogeneous, which can be observed in the phylogenetic tree in Fig. 3. This may be because of the definition of a clonal complex,

FIG 3 Whole-genome SNP-based phylogenetic tree of 154 environmental *E. coli* strains and 22 reference strains generated by kSNP3.0. The branch lengths are expressed in terms of changes per number of SNPs. The tree was visualized using FigTree (<http://tree.bio.ed.ac.uk/software/figtree>). Strain names are marked with the colors red (reference), blue (MDR), green (nonsusceptible to one or two classes), and black (susceptible). ExPEC isolates sequenced by us are marked with orange circles. Numbers (black) indicate STs/ST complexes found in five or more isolates. Phylogenetic groups (groups A, B1, B2, C, D, E, and F) are indicated near the corresponding clades. Branches are colored according to the phylogenetic group, i.e., light blue for group A, pink for group B1, light green for group B2, brown for group C, purple for group D, yellow green for group E, and orange for group F.

i.e., all STs within a clonal complex share identical alleles at 6 of the 7 loci with at least one other ST in the clonal complex (<http://eburst.mlst.net/v3/instructions/3.asp>). STC69 (clonal group A) is known to be associated with human ExPEC infections, and most strains belonging to this clonal complex are MDR (5). Actually, two of seven STC69 isolates identified in our study were classified as ExPEC and all seven isolates were MDR. STC69 is considered to be a clonal group circulating predominantly in community settings rather than health care settings (5). Our results imply that surface waters may serve as a source for infections caused by this clonal group.

TABLE 2 Distribution of sequence types and phylogenetic groups

ST or phylogenetic group	No. of isolates (%)					
	MDR		Nonsusceptible to 1 or 2 classes		Susceptible	
	ExPEC (<i>n</i> = 16)	Non-ExPEC (<i>n</i> = 50)	ExPEC (<i>n</i> = 12)	Non-ExPEC (<i>n</i> = 17)	ExPEC (<i>n</i> = 30)	Non-ExPEC (<i>n</i> = 30)
ST						
STC95	0 (0)	0 (0)	4 (33)	0 (0)	14 (47)	0 (0)
STC10	1 (6)	8 (16)	0 (0)	6 (35)	0 (0)	0 (0)
STC155	0 (0)	9 (18)	0 (0)	1 (6)	0 (0)	0 (0)
STC127	1 (6)	0 (0)	2 (17)	0 (0)	5 (17)	0 (0)
STC131	2 (13)	0 (0)	3 (25)	2 (12)	1 (3)	0 (0)
STC69	2 (13)	5 (10)	0 (0)	0 (0)	0 (0)	0 (0)
STC12	3 (19)	0 (0)	1 (8)	0 (0)	2 (7)	0 (0)
STC14	2 (13)	0 (0)	0 (0)	0 (0)	4 (13)	0 (0)
STC38	1 (6)	3 (6)	0 (0)	0 (0)	0 (0)	2 (7)
STC101	1 (6)	3 (6)	0 (0)	0 (0)	0 (0)	1 (3)
Other STs ^a	3 (19)	22 (44)	2 (17)	8 (47)	4 (13)	27 (90)
Phylogenetic group						
A	2 (13)	12 (24)	0 (0)	7 (41)	0 (0)	2 (7)
B1	1 (6)	26 (52)	0 (0)	6 (35)	0 (0)	13 (43)
B2	9 (56)	0 (0)	12 (100)	2 (12)	30 (100)	5 (17)
C	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	2 (7)
D	3 (19)	9 (18)	0 (0)	1 (6)	0 (0)	5 (17)
F	1 (6)	2 (4)	0 (0)	0 (0)	0 (0)	0 (0)
Clade V	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (3)
Not determined	0 (0)	1 (2)	0 (0)	1 (6)	0 (0)	2 (7)

^aSTs and ST complexes with fewer than five isolates are included in other STs.

STC155 was recently reported as a clonal group of animal origin that is spreading in humans and is highly drug resistant (46). Interestingly, this clonal group consisted of 10 drug-resistant isolates (9 of which were MDR) in the present study. In order to characterize this clonal group, we used Gegenees software to identify genetic regions specific to this group. However, we were not able to find genetic regions that explained the high level of antimicrobial resistance in STC155 under the stringent criteria applied. A detailed description of our Gegenees analysis is given in the supplemental material. STC155 strains have great conjugation ability and easily acquire antimicrobial resistance genes (46). Moreover, STC155 strains are able to efficiently transfer resistance to pathogenic *E. coli* strains via conjugation (46). PlasmidFinder identified plasmid replicons belonging to various incompatibility groups in 10 STC155 strains. These replicons included F (*n* = 12 [FII, *n* = 6; FIB, *n* = 6]), I1 (*n* = 2), Y (*n* = 2), and other replicons (*n* = 9) such as Col and B/O/K/Z. However, resistance genes were detected in the contigs carrying the replicons in only four cases. As mentioned above, this happened mainly because of the inability of the current sequencing technology and *de novo* assembly methods to assemble complete plasmid sequences. Some previous studies also detected STC155 strains in surface waters (13, 14). Contamination of surface waters with this highly drug-resistant clone is a concern.

We note that the definition of ExPEC used in our study is based on the presence of five virulence markers but the actual pathogenicity of strains cannot be determined merely by the presence or absence of these five markers. Therefore, an absence of the markers in MDR strains does not necessarily mean that such strains are not able to cause human extraintestinal disease.

Characteristics of ExPEC strains. As mentioned above, most of the 58 environmental ExPEC isolates belonged to phylogenetic groups B2 and D. Moreover, we found that environmental ExPEC isolates exhibited clonal distribution (Fig. 3 and Table 2), and some clinically important clonal groups, including STC95 (31%), STC127 (14%), STC12 (10%), STC14 (10%), and ST131 (10%), were prevalent in those isolates. Of those clonal groups, ST131 is of great concern because many members of this group are MDR due

to their production of CTX-M-type ESBLs and fluoroquinolone resistance (5). ST131 can be further divided into subgroups on the basis of *fimH* types (e.g., *fimH22*, *fimH30*, and *fimH41*) (10). Among these subgroups, strains carrying *fimH30* are the major cause of the global expansion of ESBL-producing ExPEC strains, and this tendency has also been observed in Japan (25). However, of six environmental ST131 isolates classified as ExPEC, five carried *fimH41* and one carried *fimH22*. Moreover, two STC131 isolates classified as non-ExPEC carried *fimH41*. This was unexpected, because strains carrying *fimH30* were found to be more prevalent among ST131 isolates than strains carrying *fimH41* or *fimH22* in clinical settings (47, 48). Although the sample size of ST131 isolates found and analyzed in the present study was small, this result suggests that environmental ST131 strains and clinical ST131 strains may differ in terms of their phylogenetic distribution, at least in our study setting. No environmental STC131 isolates produced ESBLs, only one isolate was ciprofloxacin resistant, and four of five *fimH41* isolates that were classified as ExPEC were ampicillin resistant. These results are consistent with those of a previous study that analyzed clinical ST131 isolates (48).

We found that isolates belonging to STC95 were the most common among environmental ExPEC isolates, and 14 (78%) of 18 isolates belonging to this ST complex were drug susceptible. Previous studies reported that ST95 strains are usually drug susceptible and, more importantly, are prevalent among ExPEC isolates causing community-onset infections (47, 49). Therefore, we compared the virulence gene profiles of environmental STC95 isolates with those of clinical STC95 strains and constructed a phylogenetic tree using these isolates to determine the genetic relatedness of environmental and clinical STC95 strains. We downloaded sequence data for 18 clinical STC95 strains (50–54) for comparison. Care was taken to select isolates from various studies and various types of specimens. The results are shown in Fig. 4. We found that environmental and clinical isolates were clustered together in the phylogenetic tree according to the *fimH* types, and strains belonging to the same cluster tended to carry similar (or even the same) sets of virulence genes. It was reported that clinical ST95 strains were composed of four *fimH* types, i.e., f-1, f-6, f-9, and f-47 (49), and we found that environmental STC95 isolates were composed of three (f-1, f-6, and f-47) of those four *fimH* types (it should be noted that the *fimH* classification system used by Adams-Sapper et al. [49] is different from that used in our study). These results suggested that the environmental STC95 strains obtained in our study were closely related to clinical STC95 strains. The reason why STC95 isolates are prevalent among clinical ExPEC strains although they are drug susceptible is not well understood; it may be primarily because of their enhanced capacity to cause extraintestinal infections. In light of our results, another reason may be that the probability of contacting STC95 isolates in the environment is greater than that of contacting isolates belonging to other clonal groups.

Conclusion. In the present study, we successfully characterized, by whole-genome sequencing, *E. coli* strains obtained from river water. Whole-genome analysis enabled the detection of more than 50 different resistance determinants among the isolates. We found that strains belonging to STC69, STC10, and STC155 were prevalent among MDR strains. We also demonstrated that ExPEC strains exhibited clonal distribution and that clinically important clonal groups such as STC95, STC127, STC12, STC14, and STC131 were prevalent among the isolates. Although the present study was conducted regionally, our results indicate contamination of surface waters by *E. coli* strains belonging to clinically important clonal groups, suggesting the need for further investigations of antimicrobial-resistant pathogens in surface waters.

MATERIALS AND METHODS

Bacterial isolates. *E. coli* strains from the Yamato River in Japan were isolated from 10 sites in the river between 2011 and 2013, as described previously (55). A brief description of the sampling procedure and a map of the study area are provided in the supplemental material. Isolates in which we could not detect either *adk* or *trpA* (both are *E. coli* housekeeping genes) by PCR in the previous study were removed, because those isolates were suspected to represent other species. In total, 531 isolates were subjected to antimicrobial susceptibility testing.

Antimicrobial susceptibility testing. Antimicrobial susceptibility tests were performed by microdilution using the dry plate Eiken assay (Eiken, Tokyo, Japan), following CLSI specifications. The antimicrobials used for the susceptibility testing included the following 15 categories of antimicrobial agents (with the drugs tested within each category): penicillins (ampicillin), penicillins plus β -lactamase inhibitors (amoxicillin-clavulanic acid and ampicillin-sulbactam), antipseudomonal penicillins plus β -lactamase inhibitors (piperacillin-tazobactam), broad-spectrum first-generation cephalosporins (cefazolin), extended-spectrum third- and fourth-generation cephalosporins (cefotaxime, ceftazidime, and cefepime), cephamycins (cefoxitin), carbapenems (imipenem and meropenem), monobactams (aztreonam), fluoroquinolones (ciprofloxacin), aminoglycosides (gentamicin, tobramycin, amikacin, and kanamycin), folate pathway inhibitors (trimethoprim-sulfamethoxazole), tetracyclines (tetracycline and minocycline), phenicols (chloramphenicol), polymyxins (colistin), and phosphonic acids (fosfomycin). Susceptibilities to chloramphenicol and tetracycline were also determined by the disk diffusion method. The results were interpreted according to the current ECOFF values (<http://mic.eucast.org/Eucast2/>) and 2015 CLSI criteria (56). Intermediate susceptibility to each antibiotic was considered to be resistance. MDR was defined as nonsusceptible to one or more agents in three or more antimicrobial categories, according to either ECOFF or CLSI criteria (11) (ECOFF values are equal to or less than clinical breakpoints except for those for chloramphenicol and tetracycline). Cefpodoxime and nalidixic acid were also included in the susceptibility testing, although these two antimicrobials were not used to define MDR.

Whole-genome sequencing. We excluded redundant resistant strains (i.e., strains collected from the same site at the same sampling time and exhibiting the same antimicrobial resistance and virulence gene patterns) from sequencing analysis. After exclusion, 39 ExPEC isolates identified in the previous study (3) were subjected to sequencing. In addition, 88 non-ExPEC isolates (MDR isolates, isolates nonsusceptible to one or two classes of agents, and susceptible isolates) were selected for sequencing.

DNA was extracted from each isolate by using a DNeasy blood and tissue kit (Qiagen, Hilden, Germany). A Nextera XT DNA sample preparation kit (Illumina, San Diego, CA) was used to prepare libraries for sequencing. Each library was sequenced on a MiSeq sequencer (Illumina) for 500 cycles (250-bp paired-end reads) or 600 cycles (300-bp paired-end reads), to achieve an average depth of 60 $\times$. The genomes of some strains were sequenced for 150 cycles (75-bp paired-end reads) after 600 cycles to add reads. Twenty-eight *E. coli* isolates sequenced previously by us were also included in the following analysis (NCBI/DBJ accession numbers [DRA003179](#) and [DRA003504](#)). These 28 strains are nonredundant and were selected from the 531 isolates.

In total, whole-genome sequence data for 155 nonredundant isolates were analyzed as described below (127 isolates were sequenced in the present study and 28 isolates were sequenced previously by us). Whole-genome sequencing revealed that six isolates not classified as ExPEC in the previous study actually were ExPEC (positive for two or more of *papA* and/or *papC*, *afa* and/or *dra*, *kpsMTII*, *iutA*, and *sfa* and/or *foc*) (57). This appears to have been caused by failure of PCR amplification of *kpsMTII* because of mutations in the annealing site of the reverse primer. Therefore, the actual numbers of strains whose whole genomes were analyzed were 16 isolates that were MDR and ExPEC, 50 isolates that were MDR and non-ExPEC, 12 isolates that were nonsusceptible to one or two classes of agents and ExPEC, 17 isolates that were nonsusceptible to one or two classes of agents and non-ExPEC, 30 isolates that were susceptible and ExPEC, and 30 isolates that were susceptible and non-ExPEC (also see Fig. 2 for data on the susceptibility of sequenced ExPEC strains).

Analysis of sequence data. Reads from each isolate were trimmed and assembled using CLC Genomics Workbench (CLC Bio, Aarhus, Denmark). The quality of the *de novo* assembly was ensured by using the following criteria: N50 values of at least 20,000 bp and between 4 megabases and 5.8 megabases to be assembled into contigs; these are the main criteria used for defining a good assembly in Enterobase (<http://enterobase.warwick.ac.uk>). Assembled contigs were analyzed as described below. Acquired antimicrobial resistance genes were detected using the ResFinder web server (35) and the ARG-ANNOT database (58). The *E. coli* MLST website (<http://mlst.ucc.ie/mlst/dbs/Ecoli>) and Enterobase were used for MLST. Closely related STs were grouped into ST complexes by using eBURST (<http://eburst.mlst.net>) or referring to the MLST website. The *fimH* (type 1 fimbrial adhesin gene) allele types were determined following the classification system established previously (59); SNP analysis of *fimH* can further discriminate isolates genotyped by MLST. The QRDRs of *gyrA* and *parC* and chromosomal *ampC* promoter/attenuator mutations were analyzed as described previously (27, 60). PlasmidFinder was used for identification of plasmid replicons (38).

Phylogenetic analysis was performed using kSNP3.0 software (61). The program Kchooser, which is part of the kSNP package, was used to identify the optimal kmer length and to calculate the fraction of core kmers. When Kchooser reported a value of more than 21 as an optimal k value, we chose 21 for the kmer size because k values of 19 or 21 work well for the analysis of bacterial (especially *E. coli*) genomes (62). First, we included assembled genomes from all isolates and reference genomes belonging to phylogenetic groups A, B1, B2, C, D, E, and F and cryptic clades I to V. Genomes of *Escherichia fergusonii* and *Escherichia albertii* were also included in the analysis. Reference genomes were selected from the genomes analyzed in previous studies (63, 64). The parsimony tree was estimated based on SNP loci that occurred in at least 50% of the strains. Each strain was assigned to a phylogenetic group or a cryptic clade based on its position in the phylogeny. Reference strains belonging to *E. fergusonii*, *E. albertii*, or cryptic clades and our strains found to belong to cryptic clades were then removed to increase the value of the fraction of core kmers. A parsimony tree was then estimated again as described above. It should be noted that the kSNP3 parsimony tree is a consensus of up to 100 equally parsimonious trees.

Gegenees was used to identify genomic regions specific to STC155 (65). VirulenceFinder (66) and an in-house database compiled from the Virulence Factors Database (VFDB) (67) were used for detection of virulence genes in STC95 isolates.

For all isolates, *de novo* assemblies met the required criteria. Assembly statistics, phylogenetic group, ST, resistance determinants, resistance phenotype, and plasmid replicons for each isolate are provided in Data Set S1 in the supplemental material.

Accession number(s). The genome sequence data from the present study have been deposited in the DDBJ Sequence Read Archive database (DDBJ accession number [DRA004833](https://doi.org/10.1128/AEM.02703-16)).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <https://doi.org/10.1128/AEM.02703-16>.

TEXT S1, PDF file, 0.5 MB.

DATA SET S1, XLSX file, 0.03 MB.

ACKNOWLEDGMENTS

This work was supported by Kyoto University's Global Survivability Studies program, the Sasakawa Scientific Research Grant from The Japan Science Society, and the River Fund (The River Foundation, Japan).

REFERENCES

- Ishii S, Sadowsky MJ. 2008. *Escherichia coli* in the environment: implications for water quality and human health. *Microbes Environ* 23:101–108. <https://doi.org/10.1264/jisme.2.23.101>.
- Kaper JB, Nataro JP, Mobley HL. 2004. Pathogenic *Escherichia coli*. *Nat Rev Microbiol* 2:123–140. <https://doi.org/10.1038/nrmicro818>.
- Gomi R, Matsuda T, Fujimori Y, Harada H, Matsui Y, Yoneda M. 2015. Characterization of pathogenic *Escherichia coli* in river water by simultaneous detection and sequencing of 14 virulence genes. *Environ Sci Technol* 49:6800–6807. <https://doi.org/10.1021/acs.est.5b00953>.
- Singer RS. 2015. Urinary tract infections attributed to diverse ExPEC strains in food animals: evidence and data gaps. *Front Microbiol* 6:28.
- Riley LW. 2014. Pandemic lineages of extraintestinal pathogenic *Escherichia coli*. *Clin Microbiol Infect* 20:380–390. <https://doi.org/10.1111/1469-0691.12646>.
- Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, Daszak P. 2008. Global trends in emerging infectious diseases. *Nature* 451:990–993. <https://doi.org/10.1038/nature06536>.
- Erb A, Sturmer T, Marre R, Brenner H. 2007. Prevalence of antibiotic resistance in *Escherichia coli*: overview of geographical, temporal, and methodological variations. *Eur J Clin Microbiol Infect Dis* 26:83–90. <https://doi.org/10.1007/s10096-006-0248-2>.
- Woodford N, Turtton JF, Livermore DM. 2011. Multiresistant Gram-negative bacteria: the role of high-risk clones in the dissemination of antibiotic resistance. *FEMS Microbiol Rev* 35:736–755. <https://doi.org/10.1111/j.1574-6976.2011.00268.x>.
- Manges AR, Johnson JR, Foxman B, O'Bryan TT, Fullerton KE, Riley LW. 2001. Widespread distribution of urinary tract infections caused by a multidrug-resistant *Escherichia coli* clonal group. *N Engl J Med* 345:1007–1013. <https://doi.org/10.1056/NEJMoa011265>.
- Price LB, Johnson JR, Aziz M, Clabots C, Johnston B, Tchesnokova V, Nordstrom L, Billig M, Chattopadhyay S, Stegger M, Andersen PS, Pearson T, Riddell K, Rogers P, Scholes D, Kahl B, Keim P, Sokurenko EV. 2013. The epidemic of extended-spectrum- β -lactamase-producing *Escherichia coli* ST131 is driven by a single highly pathogenic subclone, H30-Rx. *mBio* 4:e00377-13. <https://doi.org/10.1128/mBio.00377-13>.
- Magiorakos AP, Srinivasan A, Carey RB, Carmeli Y, Falagas ME, Giske CG, Harbarth S, Hindler JF, Kahlmeter G, Olsson-Liljequist B, Paterson DL, Rice LB, Stelling J, Struelens MJ, Vatopoulos A, Weber JT, Monnet DL. 2012. Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. *Clin Microbiol Infect* 18:268–281. <https://doi.org/10.1111/j.1469-0691.2011.03570.x>.
- Coleman BL, Louie M, Salvadori MI, McEwen SA, Neumann N, Sibley K, Irwin RJ, Jamieson FB, Daignault D, Majury A, Braithwaite S, Crago B, McGeer AJ. 2013. Contamination of Canadian private drinking water sources with antimicrobial resistant *Escherichia coli*. *Water Res* 47:3026–3036. <https://doi.org/10.1016/j.watres.2013.03.008>.
- Muller A, Stephan R, Nuesch-Inderbinen M. 2016. Distribution of virulence factors in ESBL-producing *Escherichia coli* isolated from the environment, livestock, food and humans. *Sci Total Environ* 541:667–672. <https://doi.org/10.1016/j.scitotenv.2015.09.135>.
- Kamruzzaman M, Shoma S, Bari SM, Ginn AN, Wiklindt AM, Partridge SR, Faruque SM, Iredell JR. 2013. Genetic diversity and antibiotic resistance in *Escherichia coli* from environmental surface water in Dhaka City, Bangladesh. *Diagn Microbiol Infect Dis* 76:222–226. <https://doi.org/10.1016/j.diagmicrobio.2013.02.016>.
- Hu J, Shi J, Chang H, Li D, Yang M, Kamagata Y. 2008. Phenotyping and genotyping of antibiotic-resistant *Escherichia coli* isolated from a natural river basin. *Environ Sci Technol* 42:3415–3420. <https://doi.org/10.1021/es7026746>.
- Jang J, Suh YS, Di DY, Unno T, Sadowsky MJ, Hur HG. 2013. Pathogenic *Escherichia coli* strains producing extended-spectrum β -lactamases in the Yeongsan River basin of South Korea. *Environ Sci Technol* 47:1128–1136. <https://doi.org/10.1021/es303577u>.
- Su HC, Ying GG, Tao R, Zhang RQ, Zhao JL, Liu YS. 2012. Class 1 and 2 integrons, *sul* resistance genes and antibiotic resistance in *Escherichia coli* isolated from Dongjiang River, South China. *Environ Pollut* 169:42–49. <https://doi.org/10.1016/j.envpol.2012.05.007>.
- Kappell AD, DeNies MS, Ahuja NH, Ledebore NA, Newton RJ, Hristova KR. 2015. Detection of multi-drug resistant *Escherichia coli* in the urban waterways of Milwaukee, WI. *Front Microbiol* 6:336. <https://doi.org/10.3389/fmicb.2015.00336>.
- Berendonk TU, Manaia CM, Merlin C, Fatta-Kassinos D, Cytryn E, Walsh F, Burgmann H, Sorum H, Norstrom M, Pons MN, Kreuzinger N, Huovinen P, Stefani S, Schwartz T, Kisand V, Baquero F, Martinez JL. 2015. Tackling antibiotic resistance: the environmental framework. *Nat Rev Microbiol* 13:310–317. <https://doi.org/10.1038/nrmicro3439>.
- Marchaim D, Chopra T, Pogue JM, Perez F, Hujer AM, Rudin S, Endimiani A, Navon-Venezia S, Hothi J, Slim J, Blunden C, Shango M, Lephart PR, Salimnia H, Reid D, Moshos J, Hafeez W, Bheemreddy S, Chen TY, Dhar S, Bonomo RA, Kaye KS. 2011. Outbreak of colistin-resistant, carbapenem-resistant *Klebsiella pneumoniae* in metropolitan Detroit, Michigan. *Antimicrob Agents Chemother* 55:593–599. <https://doi.org/10.1128/AAC.01020-10>.
- Watkinson AJ, Micalizzi GB, Graham GM, Bates JB, Costanzo SD. 2007. Antibiotic-resistant *Escherichia coli* in wastewaters, surface waters, and oysters from an urban riverine system. *Appl Environ Microbiol* 73:5667–5670. <https://doi.org/10.1128/AEM.00763-07>.
- Köhler CD, Dobrindt U. 2011. What defines extraintestinal pathogenic *Escherichia coli*? *Int J Med Microbiol* 301:642–647. <https://doi.org/10.1016/j.ijmm.2011.09.006>.
- Johnson JR, Russo TA. 2002. Extraintestinal pathogenic *Escherichia coli*: “the other bad *E. coli*.” *J Lab Clin Med* 139:155–162. <https://doi.org/10.1067/mlc.2002.121550>.

24. Robicsek A, Strahilevitz J, Jacoby GA, Macielag M, Abbanat D, Park CH, Bush K, Hooper DC. 2006. Fluoroquinolone-modifying enzyme: a new adaptation of a common aminoglycoside acetyltransferase. *Nat Med* 12:83–88. <https://doi.org/10.1038/nm1347>.
25. Matsumura Y, Johnson JR, Yamamoto M, Nagao M, Tanaka M, Takakura S, Ichihama S. 2015. CTX-M-27- and CTX-M-14-producing, ciprofloxacin-resistant *Escherichia coli* of the H30 subclonal group within ST131 drive a Japanese regional ESBL epidemic. *J Antimicrob Chemother* 70:1639–1649.
26. Jacoby GA. 2009. AmpC β -lactamases. *Clin Microbiol Rev* 22:161–182. <https://doi.org/10.1128/CMR.00036-08>.
27. Aoike N, Saga T, Sakata R, Yoshizumi A, Kimura S, Iwata M, Yoshizawa S, Sugawara Y, Ishii Y, Yamaguchi K, Tateda K. 2013. Molecular characterization of extraintestinal *Escherichia coli* isolates in Japan: relationship between sequence types and mutation patterns of quinolone resistance-determining regions analyzed by pyrosequencing. *J Clin Microbiol* 51:1692–1698. <https://doi.org/10.1128/JCM.03049-12>.
28. Matsumura Y, Yamamoto M, Nagao M, Ito Y, Takakura S, Ichihama S. 2013. Association of fluoroquinolone resistance, virulence genes, and IncF plasmids with extended-spectrum- β -lactamase-producing *Escherichia coli* sequence type 131 (ST131) and ST405 clonal groups. *Antimicrob Agents Chemother* 57:4736–4742. <https://doi.org/10.1128/AAC.00641-13>.
29. Strahilevitz J, Jacoby GA, Hooper DC, Robicsek A. 2009. Plasmid-mediated quinolone resistance: a multifaceted threat. *Clin Microbiol Rev* 22:664–689. <https://doi.org/10.1128/CMR.00016-09>.
30. Karageorgopoulos DE, Wang R, Yu XH, Falagas ME. 2012. Fosfomycin: evaluation of the published evidence on the emergence of antimicrobial resistance in Gram-negative pathogens. *J Antimicrob Chemother* 67: 255–268. <https://doi.org/10.1093/jac/ckr466>.
31. Shryock TR, Mortensen JE, Baumholtz M. 1998. The effects of macrolides on the expression of bacterial virulence mechanisms. *J Antimicrob Chemother* 41:505–512. <https://doi.org/10.1093/jac/41.5.505>.
32. Phuc Nguyen MC, Woerther PL, Bouvet M, Andremonet A, Leclercq R, Canu A. 2009. *Escherichia coli* as reservoir for macrolide resistance genes. *Emerg Infect Dis* 15:1648–1650. <https://doi.org/10.3201/eid1510.090696>.
33. Tyson GH, McDermott PF, Li C, Chen Y, Tadesse DA, Mukherjee S, Bodeis-Jones S, Kabera C, Gaines SA, Loneragan GH, Edrington TS, Torrence M, Harhay DM, Zhao S. 2015. WGS accurately predicts antimicrobial resistance in *Escherichia coli*. *J Antimicrob Chemother* 70: 2763–2769. <https://doi.org/10.1093/jac/ckv186>.
34. Perreten V, Boerlin P. 2003. A new sulfonamide resistance gene (*sul3*) in *Escherichia coli* is widespread in the pig population of Switzerland. *Antimicrob Agents Chemother* 47:1169–1172. <https://doi.org/10.1128/AAC.47.3.1169-1172.2003>.
35. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MH. 2012. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother* 67:2640–2644. <https://doi.org/10.1093/jac/dks261>.
36. Tuckman M, Petersen PJ, Howe AY, Orlowski M, Mullen S, Chan K, Bradford PA, Jones CH. 2007. Occurrence of tetracycline resistance genes among *Escherichia coli* isolates from the phase 3 clinical trials for tigecycline. *Antimicrob Agents Chemother* 51:3205–3211. <https://doi.org/10.1128/AAC.00625-07>.
37. Seputienė V, Povilonis J, Ruzauskas M, Pavilonis A, Suziedeliene E. 2010. Prevalence of trimethoprim resistance genes in *Escherichia coli* isolates of human and animal origin in Lithuania. *J Med Microbiol* 59:315–322. <https://doi.org/10.1099/jmm.0.015008-0>.
38. Carattoli A, Zankari E, Garcia-Fernandez 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>.
39. Cury J, Jove T, Touchon M, Neron B, Rocha EP. 2016. Identification and analysis of integrons and cassette arrays in bacterial genomes. *Nucleic Acids Res* 44:4539–4550. <https://doi.org/10.1093/nar/gkw319>.
40. Clermont O, Gordon D, Denamur E. 2015. Guide to the various phylogenetic classification schemes for *Escherichia coli* and the correspondence among schemes. *Microbiology* 161:980–988. <https://doi.org/10.1099/mic.0.000063>.
41. Clermont O, Gordon DM, Brisse S, Walk ST, Denamur E. 2011. Characterization of the cryptic *Escherichia* lineages: rapid identification and prevalence. *Environ Microbiol* 13:2468–2477. <https://doi.org/10.1111/j.1462-2920.2011.02519.x>.
42. Ingle DJ, Clermont O, Skurnik D, Denamur E, Walk ST, Gordon DM. 2011. Biofilm formation by and thermal niche and virulence characteristics of *Escherichia* spp. *Appl Environ Microbiol* 77:2695–2700. <https://doi.org/10.1128/AEM.02401-10>.
43. Feil EJ, Holmes EC, Bessen DE, Chan MS, Day NP, Enright MC, Goldstein R, Hood DW, Kalia A, Moore CE, Zhou J, Spratt BG. 2001. Recombination within natural populations of pathogenic bacteria: short-term empirical estimates and long-term phylogenetic consequences. *Proc Natl Acad Sci U S A* 98:182–187. <https://doi.org/10.1073/pnas.98.1.182>.
44. Hall BG. 2016. Effects of sequence diversity and recombination on the accuracy of phylogenetic trees estimated by kSNP. *Cladistics* 32:90–99. <https://doi.org/10.1111/cla.12113>.
45. Didelot X, Falush D. 2007. Inference of bacterial microevolution using multilocus sequence data. *Genetics* 175:1251–1266.
46. Skurnik D, Clermont O, Guillard T, Launay A, Danilchanka O, Pons S, Diancourt L, Lebreton F, Kadlec K, Roux D, Jiang D, Dion S, Aschard H, Denamur M, Cywes-Bentley C, Schwarz S, Tenaillon O, Andremonet A, Picard B, Mekalanos J, Brisse S, Denamur E. 2016. Emergence of antimicrobial-resistant *Escherichia coli* of animal origin spreading in humans. *Mol Biol Evol* 33:898–914. <https://doi.org/10.1093/molbev/msv280>.
47. Banerjee R, Johnston B, Lohse C, Chattopadhyay S, Tchesnokova V, Sokurenko EV, Johnson JR. 2013. The clonal distribution and diversity of extraintestinal *Escherichia coli* isolates vary according to patient characteristics. *Antimicrob Agents Chemother* 57:5912–5917. <https://doi.org/10.1128/AAC.01065-13>.
48. Johnson JR, Clermont O, Johnston B, Clabots C, Tchesnokova V, Sokurenko E, Junka AF, Maczynska B, Denamur E. 2014. Rapid and specific detection, molecular epidemiology, and experimental virulence of the O16 subgroup within *Escherichia coli* sequence type 131. *J Clin Microbiol* 52:1358–1365. <https://doi.org/10.1128/JCM.03502-13>.
49. Adams-Sapper S, Diep BA, Perdreau-Remington F, Riley LW. 2013. Clonal composition and community clustering of drug-susceptible and -resistant *Escherichia coli* isolates from bloodstream infections. *Antimicrob Agents Chemother* 57:490–497. <https://doi.org/10.1128/AAC.01025-12>.
50. Doumith M, Day M, Ciesielczuk H, Hope R, Underwood A, Reynolds R, Wain J, Livermore DM, Woodford N. 2015. Rapid identification of major *Escherichia coli* sequence types causing urinary tract and bloodstream infections. *J Clin Microbiol* 53:160–166. <https://doi.org/10.1128/JCM.02562-14>.
51. Wijetunge DS, Katani R, Kapur V, Kariyawasam S. 2015. Complete genome sequence of *Escherichia coli* strain RS218 (O18:H7:K1), associated with neonatal meningitis. *Genome Announc* 3:e00804-15. <https://doi.org/10.1128/genomeA.00804-15>.
52. Stephens CM, Skerker JM, Sekhon MS, Arkin AP, Riley LW. 2015. Complete genome sequences of four *Escherichia coli* ST95 isolates from bloodstream infections. *Genome Announc* 3:e01241-15. <https://doi.org/10.1128/genomeA.01241-15>.
53. Chen SL, Hung CS, Xu J, Reigstad CS, Magrini V, Sabo A, Blasiar D, Bieri T, Meyer RR, Ozersky P, Armstrong JR, Fulton RS, Latreille JP, Spieth J, Hooton TM, Mardis ER, Hultgren SJ, Gordon JI. 2006. Identification of genes subject to positive selection in uropathogenic strains of *Escherichia coli*: a comparative genomics approach. *Proc Natl Acad Sci U S A* 103:5977–5982. <https://doi.org/10.1073/pnas.0600938103>.
54. Salipante SJ, Roach DJ, Kitzman JO, Snyder MW, Stackhouse B, Butler-Wu SM, Lee C, Cookson BT, Shendure J. 2015. Large-scale genomic sequencing of extraintestinal pathogenic *Escherichia coli* strains. *Genome Res* 25:119–128. <https://doi.org/10.1101/gr.180190.114>.
55. Gomi R, Matsuda T, Matsui Y, Yoneda M. 2014. Fecal source tracking in water by next-generation sequencing technologies using host-specific *Escherichia coli* genetic markers. *Environ Sci Technol* 48:9616–9623. <https://doi.org/10.1021/es501944c>.
56. Clinical and Laboratory Standards Institute. 2015. Performance standards for antimicrobial susceptibility testing; 25th informational supplement. CLSI document M100-S25. Clinical and Laboratory Standards Institute, Wayne, PA.
57. Johnson JR, Murray AC, Gajewski A, Sullivan M, Snippes P, Kuskowski MA, Smith KE. 2003. Isolation and molecular characterization of nalidixic acid-resistant extraintestinal pathogenic *Escherichia coli* from retail chicken products. *Antimicrob Agents Chemother* 47:2161–2168. <https://doi.org/10.1128/AAC.47.7.2161-2168.2003>.

58. Gupta SK, Padmanabhan BR, Diene SM, Lopez-Rojas R, Kempf M, Landraud L, Rolain JM. 2014. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. *Antimicrob Agents Chemother* 58:212–220. <https://doi.org/10.1128/AAC.01310-13>.
59. Weissman SJ, Johnson JR, Tchesnokova V, Billig M, Dykhuizen D, Riddell K, Rogers P, Qin X, Butler-Wu S, Cookson BT, Fang FC, Scholes D, Chattopadhyay S, Sokurenko E. 2012. High-resolution two-locus clonal typing of extraintestinal pathogenic *Escherichia coli*. *Appl Environ Microbiol* 78:1353–1360. <https://doi.org/10.1128/AEM.06663-11>.
60. Peter-Getzlaff S, Polsfuss S, Poledica M, Hombach M, Giger J, Bottger EC, Zbinden R, Bloemberg GV. 2011. Detection of AmpC beta-lactamase in *Escherichia coli*: comparison of three phenotypic confirmation assays and genetic analysis. *J Clin Microbiol* 49:2924–2932. <https://doi.org/10.1128/JCM.00091-11>.
61. Gardner SN, Slezak T, Hall BG. 2015. kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. *Bioinformatics* 31:2877–2878. <https://doi.org/10.1093/bioinformatics/btv271>.
62. Gardner SN, Hall BG. 2013. When whole-genome alignments just won't work: kSNP v2 software for alignment-free SNP discovery and phylogenetics of hundreds of microbial genomes. *PLoS One* 8:e81760. <https://doi.org/10.1371/journal.pone.0081760>.
63. Kaas RS, Friis C, Ussery DW, Aarestrup FM. 2012. Estimating variation within the genes and inferring the phylogeny of 186 sequenced diverse *Escherichia coli* genomes. *BMC Genomics* 13:577. <https://doi.org/10.1186/1471-2164-13-577>.
64. Luo C, Walk ST, Gordon DM, Feldgarden M, Tiedje JM, Konstantinidis KT. 2011. Genome sequencing of environmental *Escherichia coli* expands understanding of the ecology and speciation of the model bacterial species. *Proc Natl Acad Sci U S A* 108:7200–7205. <https://doi.org/10.1073/pnas.1015622108>.
65. Agren J, Sundstrom A, Hafstrom T, Segerman B. 2012. Gegenees: fragmented alignment of multiple genomes for determining phylogenomic distances and genetic signatures unique for specified target groups. *PLoS One* 7:e39107. <https://doi.org/10.1371/journal.pone.0039107>.
66. Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS, Nielsen EM, Aarestrup FM. 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>.
67. Chen L, Zheng D, Liu B, Yang J, Jin Q. 2016. VFDB 2016: hierarchical and refined dataset for big data analysis: 10 years on. *Nucleic Acids Res* 44:D694–D697. <https://doi.org/10.1093/nar/gkv1239>.