Heterogeneous and Flexible Transmission of mcr-1 in Hospital-Associated Escherichia coli

Yingbo Shen  
*China Agricultural University*

Zuowei Wu  
*Iowa State University, wuzw@iastate.edu*

Yang Wang  
*China Agricultural University*

Rong Zhang  
*Zhejiang University*

Hong-Wei Zhou  
*Zhejiang University*

*See next page for additional authors*

Follow this and additional works at: [https://lib.dr.iastate.edu/vmpm_pubs](https://lib.dr.iastate.edu/vmpm_pubs)  
🔗 Part of the [Microbiology Commons](https://lib.dr.iastate.edu/vmpm_pubs), and the [Molecular Genetics Commons](https://lib.dr.iastate.edu/vmpm_pubs)  

The complete bibliographic information for this item can be found at [https://lib.dr.iastate.edu/vmpm_pubs/204](https://lib.dr.iastate.edu/vmpm_pubs/204). For information on how to cite this item, please visit [http://lib.dr.iastate.edu/howtocite.html](http://lib.dr.iastate.edu/howtocite.html).
Heterogeneous and Flexible Transmission of mcr-1 in Hospital-Associated Escherichia coli

Abstract
The recent emergence of a transferable colistin resistance mechanism, MCR-1, has gained global attention because of its threat to clinical treatment of infections caused by multidrug-resistant Gram-negative bacteria. However, the possible transmission route of mcr-1 among Enterobacteriaceae species in clinical settings is largely unknown. Here, we present a comprehensive genomic analysis of Escherichia coli isolates collected in a hospital in Hangzhou, China. We found that mcr-1-carrying isolates from clinical infections and feces of inpatients and healthy volunteers were genetically diverse and were not closely related phylogenetically, suggesting that clonal expansion is not involved in the spread of mcr-1. The mcr-1 gene was found on either chromosomes or plasmids, but in most of the E. coli isolates, mcr-1 was carried on plasmids. The genetic context of the plasmids showed considerable diversity as evidenced by the different functional insertion sequence (IS) elements, toxin-antitoxin (TA) systems, heavy metal resistance determinants, and Rep proteins of broad-host-range plasmids. Additionally, the genomic analysis revealed nosocomial transmission of mcr-1 and the coexistence of mcr-1 with other genes encoding β-lactamases and fluoroquinolone resistance in the E. coli isolates. These findings indicate that mcr-1 is heterogeneously disseminated in both commensal and pathogenic strains of E. coli, suggest the high flexibility of this gene in its association with diverse genetic backgrounds of the hosts, and provide new insights into the genome epidemiology of mcr-1 among hospital-associated E. coli strains.

Keywords
E. coli, genetic diversity, mcr-1, population genomics, transmission

Disciplines
Genetics and Genomics | Microbiology | Molecular Genetics

Comments

Creative Commons License
This work is licensed under a Creative Commons Attribution 4.0 License.

Authors
Yingbo Shen, Zuowei Wu, Yang Wang, Rong Zhang, Hong-Wei Zhou, Shaolin Wang, Lei Lei, Mei Li, Jiachang Cai, Jonathan Tyrrell, Guo-Bao Tian, Congming Wu, Qijing Zhang, Jianzhong Shen, Timothy R. Walsh, and Zhangqi Shen
Heterogeneous and Flexible Transmission of \textit{mcr-1} in Hospital-Associated \textit{Escherichia coli}

Yingbo Shen, Zuowei Wu, Yang Wang, Rong Zhang, Hong-Wei Zhou, Shaolin Wang, Lei Lei, Mei Li, Jiachang Cai, Jonathan Tyrrell, Guo-Bao Tian, Congming Wu, Qijing Zhang, Jianzhong Shen, Timothy R. Walsh, Zhangqi Shen

\textbf{ABSTRACT} The recent emergence of a transferable colistin resistance mechanism, MCR-1, has gained global attention because of its threat to clinical treatment of infections caused by multidrug-resistant Gram-negative bacteria. However, the possible transmission route of \textit{mcr-1} among \textit{Enterobacteriaceae} species in clinical settings is largely unknown. Here, we present a comprehensive genomic analysis of \textit{Escherichia coli} isolates collected in a hospital in Hangzhou, China. We found that \textit{mcr-1}-carrying isolates from clinical infections and feces of inpatients and healthy volunteers were genetically diverse and were not closely related phylogenetically, suggesting that clonal expansion is not involved in the spread of \textit{mcr-1}. The \textit{mcr-1} gene was found on either chromosomes or plasmids, but in most of the \textit{E. coli} isolates, \textit{mcr-1} was carried on plasmids. The genetic context of the plasmids showed considerable diversity as evidenced by the different functional insertion sequence (IS) elements, toxin-antitoxin (TA) systems, heavy metal resistance determinants, and Rep proteins of broad-host-range plasmids. Additionally, the genomic analysis revealed nosocomial transmission of \textit{mcr-1} and the coexistence of \textit{mcr-1} with other genes encoding \beta-lactamases and fluoroquinolone resistance in the \textit{E. coli} isolates. These findings indicate that \textit{mcr-1} is heterogeneously disseminated in both commensal and pathogenic strains of \textit{E. coli}, suggest the high flexibility of this gene in its association with diverse genetic backgrounds of the hosts, and provide new insights into the genome epidemiology of \textit{mcr-1} among hospital-associated \textit{E. coli} strains.

\textbf{IMPORTANCE} Colistin represents one of the very few available drugs for treating infections caused by extensively multidrug-resistant Gram-negative bacteria. The recently emergent \textit{mcr-1} colistin resistance gene threatens the clinical utility of colistin and has gained global attention. How \textit{mcr-1} spreads in hospital settings remains unknown and was investigated by whole-genome sequencing of \textit{mcr-1}-carrying \textit{Escherichia coli} in this study. The findings revealed extraordinary flexibility of \textit{mcr-1} in its spread among genetically diverse \textit{E. coli} hosts and plasmids, nosocomial transmission of \textit{mcr-1}-carrying \textit{E. coli}, and the continuous emergence of novel Inc types of plasmids carrying \textit{mcr-1} and new \textit{mcr-1} variants. Additionally, \textit{mcr-1} was found to be frequently associated with other genes encoding \beta-lactams and fluoroquinolone resistance. These findings provide important information on the transmission and epidemiology of \textit{mcr-1} and are of significant public health importance as the in-
formation is expected to facilitate the control of this significant antibiotic resistance threat.

KEYWORDS E. coli, genetic diversity, mcr-1, population genomics, transmission

The relentless increase in the populations of multidrug-resistant (MDR) and extensively drug-resistant (XDR) Gram-negative bacterial strains is worrying, not least because we apparently have no new clinical options. Moreover, the antibiotic pipeline is bereft of novel entities to potentially cover MDR and XDR Gram-negative infections (1–4). Recent global attention has focused on the plight of our entry into the “post-antibiotic era” in the face of the rapid dissemination of carbapenem-resistant mechanisms (NDM-1, KPC, and OXA-48/181/232) and the realization of the very limited number of antibiotics, e.g., colistin, that we have left to treat serious infections (5–7). Until recently, colistin resistance was observed to be mediated by chromosomal mutations only and commonly occurred in Klebsiella pneumoniae and Pseudomonas aeruginosa but rarely in Escherichia coli (8). However, the first transferable colistin resistance mechanism, termed mcr-1, was recently reported in Enterobacteriaceae from both food-producing animals and human origins, in particular, E. coli (9). Perhaps of greater concern is the coexistence of mcr-1 and carbapenem resistance genes, such as blaNDM-5/9 and blaKPC-2 recently identified in E. coli from human infections and poultry production (10–13), as the common occurrence of the mcr-1 gene in carbapenem-resistant Enterobacteriaceae (CRE) would seriously compromise current treatment options not just in China but globally.

The unprecedented global increase in the populations of CRE, and now of mcr-1-positive Enterobacteriaceae (MCRPE), has placed further pressure on drug discovery programs to produce novel antimicrobials. It is still unclear what drives CRE and MCRPE, and although the increase in antibiotic consumption (e.g., the use of carbapenems) has been attributed to this increase, the remarkable plasticity and fluidity of DNA structures in Gram-negative bacteria have made a significant contribution. In Enterobacteriaceae, this horizontal gene transfer is fueled by a potent cocktail of plasmids, transposons, insertion sequence (IS) elements, and insertion sequence common region (ISCR) elements (14, 15). For example, the blaNDM-1 gene can be found in considerably greater numbers of bacteria than its KPC counterpart—in part, as a result of the diversity of plasmids it is associated with (16). The immediate genetic context surrounding the blaNDM-1 gene is also remarkably heterogeneous and has contributed to its translocation between chromosome and plasmid, and vice versa, and between plasmids (16).

Due to its global significance, many studies have been reported on investigating the prevalence of mcr-1 and have characterized its genetic environments in Enterobacteriaceae. To date, more than 70 completed sequences of plasmids carrying mcr-1 have been deposited into the GenBank database and their data show that they are relatively narrow in range and contain few other antibiotic resistance genes (11, 17–27). Thus far, many whole-genome sequences have been retrospectively searched for mcr-1; however, the majority of the reports focus only on the gene or plasmid and few have analyzed its associated bacterial hosts (18, 28). Moreover, data on the possible transmission routes of mcr-1 among MCRPE are largely lacking.

In order to aid understanding of the prevalence and outcomes of the presence of MCRPE in patients as well as in healthy adults, two studies were recently published on mcr-1-positive isolates and their impacts on nosocomial infections (29, 30). Statistical data suggest that mcr-1-positive E. coli (MCRPEC) infections were found to be associated with male sex, immunosuppression, and antibiotic usage (30). Multiple studies have also shown that E. coli is the most frequently observed Enterobacteriaceae species carrying the mcr-1 gene and that the gene can be transferred to other Enterobacteriaceae species from MCRPEC at high frequencies (9, 29, 30). However, despite those studies, there is a marked paucity of genetic data on MCRPE in hospitals that can help researchers to understand their circulation and thus their potential impact on infection control policies. Here, we present an extensive whole-genome analysis of 80 E. coli
strains isolated from both clinical samples and fecal samples of patients and healthy human volunteers in one hospital in Hangzhou, Zhejiang Province, in China.

RESULTS

Overview of mcr-1-positive E. coli. During our previous epidemiological and clinical study, we obtained a considerable number of MCRPEC isolates from patients and healthy adults in Hangzhou, Zhejiang, China (30). Here, we performed whole-genome analysis of 80 MCRPEC strains to address the possible dissemination of the mcr-1 gene (see Table S1 in the supplemental material). Briefly, 36 MCRPEC isolates were derived from nosocomial infections (urinary tract, surgical wounds, respiratory tract, etc.) of inpatients \((n = 2,577)\), while 27 were from fecal samples of inpatients \((n = 1,028)\) and 17 were from fecal samples of healthy volunteers \((n = 2,909)\) collected in 2015. Detailed clinical information on these 80 mcr-1-carrying E. coli isolates is presented in Table S2.

Genome sequencing of the 80 mcr-1-positive E. coli isolates. At least 100× coverage of raw reads from Illumina sequencing was obtained for each isolate. The draft genomes were assembled de novo using CLC Genomics Workbench (version 8.5). The number of contigs ranged from 51 to 232, while the N50 of contigs ranged from 43 kb to 521 kb (Table S1) for the isolates assembled by CLC Genomics Workbench (version 8.5). Since the majority of the mcr-1 genes were located on the plasmids, we further used the plasmidSPAdes program to optimize the assembly of plasmids. The mcr-1-containing contigs for each isolate were extracted from the two assemblies, and the longer contig was used to determine the genetic context of mcr-1. The lengths of mcr-1-carrying contigs generated by Illumina sequencing ranged from 2,388 to 141,207 bp (Table S1). Due to the shortness of the reads generated by Illumina sequencing and the high number of insertion elements, the assembled mcr-1-carrying contigs for nine isolates were short (2.3 to 3.8 kb). Thus, these isolates were resequenced by single-molecule real-time (SMRT) sequencing to generate complete chromosomes and plasmids (Table S1 and S3). A further six isolates were also subjected to SMRT sequencing.

Genomic epidemiology of the mcr-1-carrying E. coli isolates. The genomic and epidemiological relationships among all MCRPEC isolates were investigated (Fig. 1). Classification of strains to phylogenetic subgroups (31, 32) and of sequence types using multilocus sequence typing (MLST) (33) was performed through in silico analysis (Fig. 1). The core genome phylogenetic tree indicates 3 main clusters, but MLST analysis revealed that the 80 MCRPEC isolates were significantly diverse: 76 isolates were assigned to 46 known MLST types, and 4 other isolates (ST6399, ST6404, ST6405, and ST6406) possessed novel STs (Fig. 1). The most prevalent ST of isolates harboring mcr-1 gene was ST10, accounting for 12.5% of all isolates. This ST is also commonly observed in extended-spectrum-β-lactamase (ESBL)-carrying E. coli isolates of both human and animal origins (34,35). The classification of phylogenetic subgroups demonstrated that the 80 isolates were distributed throughout the four phylogroups (A, B1, B2, and D) (32), but the majority fell within the A and B1 groups, with 51.3% \((n = 41)\) in the A group and 33.8% \((n = 27)\) in the B1 group. Only 6.3% \((n = 5)\), and 8.8% \((n = 7)\) of the strains belonged to group B2 and group D, respectively (Fig. 1). Analyzing the etiology of group A and B1 isolates, we found that intestinal and extraintestinal isolates were extensively disseminated throughout each phylogenetic group, without significant enrichment in either of the phylogroups \(P > 0.05\) [Fisher’s exact test]. These results indicated that no clear phylogenomic division exists between intestinal strains and extraintestinal strains as observed in the disparate sequence types (such as ST10, ST48, ST156, etc.) (Fig. 1). Similar distributions were observed in groups B2 and D, although each group contained only a limited number of isolates.

VFs and antibiotic resistance profiles of mcr-1-carrying isolates. The virulence factors (VFs) in E. coli play an important role in conferring selective advantages and defining pathogenicity profiles. We therefore cataloged the known VFs, including genes associated with adherence, autotransporters, invasion, iron uptake, toxins, and secre-
FIG 1 Genomic analysis of mcr-1-carrying E. coli isolates in a single hospital system in Zhejiang Province, China. A maximum-likelihood phylogenetic tree was constructed using the core genome SNPs and midpoint rooted. Sources of the isolates are indicated by different colors for strain identification (ID) plus MLST (red, infectious sample of inpatient; black, feces of healthy volunteer). E. coli phylogroups are denoted by colored strips, and the branches of the tree are colored in correspondence to the coloring of the strips. The location of mcr-1 on a plasmid or chromosome (dark red) and the presence or absence of virulence genes (green) and antibiotic resistance genes (orange) are denoted by filled and empty squares, respectively. Only categories of the virulence genes and 9 clinically important antibiotic resistance genes are shown. Details of the genes in each category are given in Table S4 and S5.
tion systems, as well as the antibiotic resistance phenotypes. The 12 isolates from phylogroups B2 and D possessed the greatest number of VFs, while 68 isolates of phylogroup A and B strains displayed low prevalence of VFs (Fig. 1; see also Table S4). However, no significant difference was observed in the frequency of VFs between the intestinal and extraintestinal isolates. Furthermore, no VFs related to intestinal pathogenic E. coli could be found, which is consistent with the clinical observation of the patients. Several genes encoding fimbriae were commonly present among the isolates, including those encoding E. coli common pilus (ECP), EaeH, and type 1 fimbriae; however, colonization factor antigen I (CFA/I) fimbriae were less frequently associated with phylogroup A. Types P and S fimbriae, related to extraintestinal infection, were found in phylogroups B2 and D (Fig. 1). Antibiotic resistance profiling showed that all strains were resistant to multiple categories of drugs, including colistin. Two isolates, ZJ134 (from sputum of an inpatient) and ZJ33 (from feces of an inpatient), were resistant to 11 classes of antimicrobials (Table S2); however, the resistance profiles of the isolates were not specific to any phylogenetic group. A total of 58.8% (47/80) of the isolates carried both mcr-1 and ESBL genes (Table S5). Of particular note, one isolate, ZJ487 (ST90), from intra-abdominal fluid contained not only mcr-1 but also blaCTX-M-1, blaNDM-1, blaCTX-M-55, and blaSHV-12. E. coli ST131 is the most prevalent sequence type associated with extraintestinal infections (36–38), and we identified two mcr-1-positive ST131 isolates of the serotype of H4:O25. One ST131 isolate, ZJ26, from the fecal sample of a healthy volunteer, carried VF genes iroN, iha, gad, iss, cma, tsh, and fimH22 and antibiotic resistance genes blACTX-M-55 and mcr-1 and the qpxAB plasmid-mediated quinolone resistance gene. The other ST131 isolate, ZJ71, from the fecal sample of a healthy volunteer, possessed fluoroquinolone resistance alleles gyrA1AB and parC1aAB, blACTX-M-1a, blTEM-1B, and VF genes sat, iss, iha, and senB; the fimH30-gyrA1AB-parC1aAB allelic profile of that isolate was considered to show that it was the more pandemic of the two ST131 isolates (39).

mcr-1 is mainly plasmid mediated. To classify the locations of mcr-1 in the genomes, mcr-1-containing contigs were extracted from the complete genomes of 15 isolates subjected to SMRT sequencing and the draft genomes of 65 isolates sequenced by Illumina. The results of analysis of the complete genomes of 15 isolates suggested that mcr-1 is located on the chromosome in 6 isolates and on plasmids in 9 isolates (Fig. 2; see also Table S1 and S3). Interestingly, isolates ZJ1432 and ZJ859 contain two copies of mcr-1, which were located on distinct plasmids (Fig. 2). mcr-1 was carried on an ~34-kb IncX4 plasmid and an ~62-kb IncI2 plasmid in ZJ1432, while mcr-1 was carried on an ~68-kb IncI2 plasmid and an ~82-kb IncFIB plasmid in ZJ859. To determine the locations of mcr-1 in the rest of the 65 isolates sequenced by Illumina, mcr-1-containing contigs from the 65 isolates were searched against the PlasmidFinder database (40) to define their plasmid types, and the whole draft genomes of each of the isolates were subjected to a BLAST search against the complete mcr-1-carrying plasmids deposited in GenBank to find the closest reference. PlasmidFinder detected that 54 of the 65 mcr-1-positive contigs also carried plasmid replicons, suggesting that their mcr-1 genes are plasmid-borne genes. The other 11 isolates did not possess any known plasmid replicons within the mcr-1-positive contigs. When the complete set of contigs for each isolate was mapped to reference plasmids, 10 of 11 isolates mapped more than 95% to their closest reference plasmids (Fig. 2; see also Table S1). Replicons of the 10 isolates were reexamined using the complete set of contigs of each genome. The results showed that all 10 isolates contain the same replicons as the reference plasmids (see Fig. S1 to S4 at https://doi.org/10.6084/m9.figshare.6281579.v1), but the replicons were separated from mcr-1-containing contigs due to the incomplete assembly. The only exception was the mcr-1 contig (~19.7 kb) from isolate ZJ3951, which did not contain any plasmid replicons and did not map sufficiently with any known plasmids. However, that contig is highly similar to homologous regions of the E. coli chromosome available in the GenBank database (e.g., CP018962.1 and CP018103.1), indicating that mcr-1 in ZJ3951...
is located on the chromosome (see Fig. S5 at https://doi.org/10.6084/m9.figshare.6281579.v1). As expected, the locations of \textit{mcr}-1 based on the whole-genome sequence analysis were highly consistent with the S1 pulsed-field gel electrophoresis (PFGE) results (see Fig. S6 at https://doi.org/10.6084/m9.figshare.6281579.v1), represented as chromosome-borne or plasmid-borne. In total, \textit{mcr}-1 was found to be chromosome-borne in 7 isolates and plasmid-borne in 73 isolates. As indicated in the NCBI database, IncI2 (n/H11005 29, 38.7\%) and IncX4 (n/H11005 28, 37.3\%) were the two main plasmid types carrying \textit{mcr}-1 gene (Fig. 2). A 97-kb p0111 plasmid was first found to carry the \textit{mcr}-1 gene by SMRT sequencing in two isolates, ZJ275 and ZJ276.

**Genetic context of \textit{mcr}-1 in the 80 \textit{E. coli} isolates.** To characterize the mobile genetic elements and antibiotic resistance genes in the immediate proximity of \textit{mcr}-1, we extracted the \textit{mcr}-1-bearing contigs from 65 draft genomes and 15 complete genomes. Of the 65 contigs from draft genomes, 63 ranged from 12.1 kb to 141.2 kb in size; the remaining two were ~6.7 kb and ~8.3 kb in ZJ1682 and ZJ878, respectively (Table S1). We classified the genetic context into the following nine groups: an IncI2-related group (n = 29), an IncX4-related group (n = 28), an IncHI2-related group (n = 2), an IncP-related group (n = 2), an Incp0111-related group (n = 2), an IncY-related group (n = 1), an IncFIB-related group (n = 1), a non-Inc-related group (n = 10), and a chromosome-related group (n = 7). Overall, considerable diversity was found in the immediate environment of \textit{mcr}-1. Compared to the previous reports, several novel genetic environments of \textit{mcr}-1 were found in this study, especially among those located on the chromosomes (Fig. 3).

**IncI2-related contigs.** The contigs from 29 isolates were IncI2-related (Fig. 2; see also Fig. S7 at https://doi.org/10.6084/m9.figshare.6281579.v1). The flanking sequences of \textit{mcr}-1 in all 29 isolates were very similar to sequences in plasmid pHNSHP45 (IncI2; KP347127.1), the first reported \textit{mcr}-1-carrying plasmid in \textit{E. coli} of swine origin in China. The IS\textit{Apl}1 insertion sequence is located immediately upstream of \textit{mcr}-1 in pHNSHP45. However, only three of our isolates (ZJ487, ZJ859, and ZJ3903) contained IS\textit{Apl}1 upstream of \textit{mcr}-1, while the other 26 isolates showed no evidence of IS\textit{Apl}1 adjacent to \textit{mcr}-1 (Fig. 3A; see also Fig. S7 at https://doi.org/10.6084/m9.figshare.6281579.v1).
Among the 26 isolates, 18 did not contain any insertion sequence and 8 contained IS elements other than IS\textit{Apl1} (Fig. 3A). IS\textit{Sen6}, initially found in \textit{Salmonella enterica} (NC_006511), was present in \textit{mcr-1}-carrying contigs in four isolates (ZJ1336, ZJ315, ZJ9, and ZJ257). IS\textit{Kpn26}, initially found in \textit{K. pneumoniae} (NC_016845), was detected in \textit{mcr-1}-positive contigs in ZJ148, while IS\textit{Ecp1} and IS\textit{1294}, initially found in \textit{E. coli}, were identified in \textit{mcr-1}-positive contigs in ZJ33, ZJ89, and ZJ1635. Even though multiple ESBL genes were detected in isolate ZJ487, only \textit{bla\textsubscript{CTX-M-1}} was found to be located on the \textit{mcr-1}-carrying plasmid shown by SMRT sequencing. \textit{bla\textsubscript{NDM-1}} and \textit{bla\textsubscript{SHV-12}} were found to be located on an ~54-kb plasmid, and \textit{bla\textsubscript{CTX-M-55}} and \textit{bla\textsubscript{OXA-1}} were found to be located on another, ~129-kb plasmid. In isolate ZJ33, \textit{bla\textsubscript{CTX-M-64}} flanked by IS\textit{Ecp1} was shown to be located in the same \textit{mcr-1}-carrying contig. Isolate ZJ859 harbored two IS\textit{609} copies and one IS\textit{2} copy in a \textit{mcr-1}-carrying contig.

**IncX4-related contigs.** Twenty-eight contigs were associated with IncX4 (Fig. 2; see also Fig. S8 at https://doi.org/10.6084/m9.figshare.6281579.v1). The overall genetic
contexts of these contigs showed similarity to plasmid pECGD-8-33 (IncX4; KX254343.1), and 26 of them did not show any evidence of insertion sequences (Fig. 3B; see also Fig. S8 at https://doi.org/10.6084/m9.figshare.6281579.v1). However, a truncated insertion sequence, ISKpn26, was found upstream of mcr-1 in ZJ1653, and IS26 was found ~2 kb upstream of mcr-1 in ZJ812. It is known that IS26 is commonly associated with resistance genes in E. coli and other species (Fig. 3B) (41). However, no resistance genes were found on the mcr-1-carrying contigs of ZJ1653 or on the other IncX4-related contigs.

**IncHI2-related contigs.** Two mcr-1-positive contigs from isolates ZJ675 and ZJ650 were found to be IncHI2-related (Fig. 2; see also Fig. S9 at https://doi.org/10.6084/m9.figshare.6281579.v1), and the genetic context was found to be similar to that of pHNSHP45-2 (IncHI2; KU341381.1). The mcr-1-carrying contig in ZJ650 was 107,130 bp and did not harbor any IS elements or other antibiotic resistance genes. The mcr-1-positive contig in ZJ675 (141,207 bp) contains tellurium resistance system genes, i.e., a ter (tellurium resistance [Ter] operon (terZABCDEF), conferring resistance to tellurite (42). In addition, IS150 and Tn3 were identified in the mcr-1-positive contig of ZJ675 (Fig. 3D; see also Fig. S9 at https://doi.org/10.6084/m9.figshare.6281579.v1).

**IncP-related contigs.** The genetic contexts of mcr-1-carrying contigs from isolates ZJ800 and ZJ623 are IncP-related and resemble DNA regions from plasmid pMCR_1511 (IncP; KX377410.1) (Fig. 2; see also Fig. S10 at https://doi.org/10.6084/m9.figshare.6281579.v1). ISAp1 is missing in the two mcr-1-carrying contigs (Fig. 3C); instead, IS26 is distally present with mcr-1. No other antibiotic resistance genes were found in these two contigs.

![FIG 3 (Continued)](image)
Incp0111-, IncY-, and IncFIB-related contigs. The complete mcr-1 plasmids in ZJ275 and ZJ276 possess the Incp0111 replicon (see Fig. S11 at https://doi.org/10.6084/m9.figshare.6281579.v1), while the complete mcr-1-positive plasmid in ZJ338 contains the IncY replicon (see Fig. S12 at https://doi.org/10.6084/m9.figshare.6281579.v1). The genetic contexts of Incp0111 and IncY are very similar, and mcr-1 was flanked by ISAp1l at both ends, represented as ISAp1l–mcr-1–hp–ISAp1l (Fig. 3F; see also Fig. S11 at https://doi.org/10.6084/m9.figshare.6281579.v1), with the two ISAp1l elements in the same orientation. Inverse PCR using primers located within mcr-1 was performed to detect the small circular form of this region in ZJ275 and ZJ276. PCR produced a 3,696-bp amplicon containing mcr-1, the region between mcr-1 and two ISAp1l genes, and one intact ISAp1l element (data not shown), suggesting that recombination between the two ISAp1l copies may occur and may facilitate the transmission of mcr-1. One complete plasmid (81,972 bp) of ZJ859 possesses the IncFIB replicon (see Fig. S13 at https://doi.org/10.6084/m9.figshare.6281579.v1), in which mcr-1 is flanked by ISAp1l and IS26 upstream and IS26 downstream (Fig. 3; see also Fig. S14 at https://doi.org/10.6084/m9.figshare.6281579.v1). Furthermore, genes corresponding to Tn1, IS1A, another two IS26 elements, and a type IV secretion system were present in the plasmid (Fig. 3F). This plasmid also contains two additional antibiotic resistance genes, blatem-1 and tetM, which were associated with Tn3 and proximal to mcr-1.

Non-Inc-related contigs. Ten mcr-1-positive contigs did not have replicon genes due to incomplete assembly (Fig. 2). As described previously, the possible Inc types of these contigs carrying plasmids were predicted through mapping the contigs to the reference plasmids (see Fig. S1 to S4 at https://doi.org/10.6084/m9.figshare.6281579.v1). The mcr-1-positive contigs of ZJ53, ZJ54, ZJ116, ZJ119, and ZJ878 were similar to those of pHNSHP45 (IncI2) (see Fig. S1 at https://doi.org/10.6084/m9.figshare.6281579.v1); the mcr-1-positive contig of ZJ116 contains a copy of ISSen6. In isolate ZJ2922, ISAp1l–mcr-1–hp was located downstream of a 7,689-bp segment that exhibited 92.5% nucleotide identity to the corresponding region of plasmid pCFSAN002050 from Salmonella enterica (GenBank accession number CP006056.1) (Fig. 3G). In isolates ZJ134, ZJ346, ZJ299, and ZJ675, the mcr-1-positive contigs contain the tellurium resistance system, which is commonly carried by IncHI2 or IncHI1 plasmids (43). Moreover, a Tn7-like transposon element and two resistance systems for heavy metals, the pco (COPR) operon (pcoEADBCDRSE) for copper resistance and the sil (sir) operon (sileSRCBAP) for silver resistance, were identified upstream of ISAp1l–mcr-1–hp in isolate ZJ299, which is similar to the corresponding region of IncHI2 plasmid pAPEC-O1-R (DQS17526) from E. coli isolated from turkey originating from United States (44) (Fig. 3G).

Chromosome-borne mcr-1. mcr-1 was located on the chromosome in seven isolates; six of the isolates were subjected to SMRT sequencing, and the remaining isolate, ZJ3951, was subjected to Illumina sequencing (Fig. 3E; see also Fig. S15 at https://doi.org/10.6084/m9.figshare.6281579.v1). The mcr-1-carrying contig of ZJ3951 is ~19.7 kb in size and similar to the chromosome of E. coli G27 from Belgium (accession number LOOY01000001.1) (Fig. 3E). In six isolates, ZJ40, ZJ167, ZJ250, ZJ393, ZJ1323, and ZJ1794, mcr-1 was flanked by two ISAp1l elements, forming a unit of ISAp1l–mcr-1–hp–ISAp1l as observed in plasmids from ZJ275 and ZJ276. Interestingly, mcr-1 was inserted in the same location in only two isolates, ZJ1323 and ZJ250, which belong to ST4014, and was located in the quite different regions in the other five isolates. All five isolates belong to different ST clades, and the integration sites of mcr-1 on the chromosome were distinct (Fig. 3E), indicating that the mobile genetic unit of “ISAp1l–mcr-1–hp–ISAp1l” has been recently disseminated among E. coli isolates.

mcr-1 allelic variations. Here, 10 MCR-1 variants (mcr-1.2 to mcr-1.11) were deposited in NCBI, and in this study novel mcr-1 variants were detected in six isolates (ZJ10, ZJ2431, ZJ299, ZJ3920, ZJ346, and ZJ432) from various sources (Table S6). The allelic variants from isolates ZJ10, ZJ2431, and ZJ346 exhibited synonymous mutations at positions 18 (T→C), 27 (C→T), and 552 (C→T), respectively. Moreover, ZJ299 and ZJ3920 each harbored a single nucleotide mutation in mcr-1, resulting in nonsynon-
mous mutations at positions 145 (G→A, Gly→Ser), and 1423 (G→A, Val→Ile), respectively. Finally, isolate ZJ1432, which had two mcr-1 copies in two different plasmids, contained a synonymous mutation at site 27 (C→T) in the mcr-1 copy on the IncX4 plasmid. All the mutations were confirmed by Sanger sequencing (data not shown).

DISCUSSION

In this study, 80 mcr-1-positive E. coli isolates of intestinal or extraintestinal sources from a single hospital in China were analyzed by a combination of Illumina and SMRT sequencing. We observed significant diversity not only in the genetic context of mcr-1 but particularly in the host E. coli. The mcr-1-positive E. coli isolates are distributed throughout the four phylogroups (A, B1, B2, and D) (32) and represented by 50 ST clades, indicating a lack of clonal spread and implying no outbreak of mcr-1-positive E. coli in this hospital (Fig. 1). In a recent study conducted by another Chinese group, MCRPEC isolates of fecal sources were also found to be genetically diverse (45). However, a case of possible nosocomial transmission of mcr-1-positive E. coli was observed. The complete genomes of ZJ275 and ZJ276 revealed that the two isolates contain identical chromosomes and plasmids, with only 32 single nucleotide variants (SNVs) (4 single nucleotide polymorphisms [SNPs] and 28 indels) in the chromosome and 2 SNVs in the mcr-1-positive plasmid. It should be noted that the mcr-1 gene was located on p0111 plasmids, which were specific to these two strains. Furthermore, ZJ275 and ZJ276 were isolated from the puncture fluid of a 76-year-old man and from the drainage fluid of a 38-year-old woman. The patients were kept in the same surgical intensive care unit during the same period and therefore, the possibility exists of the drainage fluid of a 38-year-old woman. The patients were kept in the same surgical intensive care unit during the same period and therefore, the possibility exists of nosocomial transmission of mcr-1-positive E. coli, highlighting the importance of hospital-acquired infections and hospital infection control programs (46).

Previous epidemiological studies indicated that groups A and B1 are usually composed of commensal strains or intestinal pathogenic strains (47) and groups B2 and D are associated with extraintestinal pathogenic strains (36, 48). However, in this study, no clear phylogenomic divisions were observed between intestinal strains and extraintestinal strains. Interestingly, examination of 14 extraintestinal isolates from bloodstream (n = 2) and urinary tract (n = 12) showed that only two isolates belonged to groups B2 and D (Fig. 1). Furthermore, analysis of virulence factors did not show any correlation between intestinal isolates and extraintestinal isolates (Fig. 1; see also Table S4 in the supplemental material). Although the data set is limited, this observation suggests that the intestinal strains may potentially cause systemic infections and serve as a source of nosocomial infections. In hospital settings, many patients are immunocompromised, have indwelling urinary catheters, and are exposed to numerous antimicrobials (Table S2), which might promote extraintestinal infection by antibiotic-resistant E. coli strains that are not normally considered extraintestinal pathogenic E. coli strains.

Acquisition of antibiotic resistance determinants through horizontal gene transfer is highly problematic. The presence of antibiotic resistance gene-linked transferable elements facilitates their spread among different clones and different bacterial species (49). mcr-1 can be found in various plasmids and in different locations on the chromosomes of E. coli. The plasmids were classified into 8 Inc types by bioinformatics analysis, and the IncI2 (38.7%) and IncN4 (37.3%) plasmids were found to be the most prevalent types among the isolates, consistent with previous reports (21, 50, 51). IncI2 mcr-1-carrying plasmids were previously found to be dominant in E. coli-associated bloodstream infections (29). However, our study data indicate that neither IncI2 plasmids nor IncN4 plasmids were predominantly associated with intestinal or extraintestinal sources, patients (6/27 IncI2, 11/27 IncN4), or volunteers (8/17 IncI2, 5/17 IncN4). Interestingly, the genetic contexts of mcr-1 were found to be highly divergent, and yet the majority of mcr-1 genes were not adjacent to IS elements (17 were flanked with IS elements [Fig. 3]). ISAp1 was the most common IS element and was found to be adjacent to mcr-1 at one end or both ends in 15 isolates. The presence of ISAp1 at the both ends of mcr-1 makes the ISAp1–mcr-1–hp–ISAp1 unit potentially active, as was evidenced by the circular form shown by inverse PCR, which can integrate into various
locations on the chromosome and/or plasmids in a manner consistent with previous reports (52, 53). Furthermore, ISKpn26 was adjacent to mcr-1 in isolates ZJ148 and ZJ1653, and several other IS elements such as IS2, IS609, IS683, IS1294, IS150, and ISSen6 were found in the mcr-1-positive contigs but not in locations immediately proximal to mcr-1. Whether these IS elements could contribute to the mobilization of mcr-1 remains unknown, but their presence would enhance the transmission of mcr-1 and facilitate its close association with other antibiotic resistance genes as seen in isolates ZJ134, ZJ33, ZJ487, ZJ26, and ZJ71 (41, 54, 55).

In this study, a strategy combining Illumina sequencing and SMRT sequencing was applied to decipher the genetic diversity of mcr-1-positive E. coli isolates. Although we successfully decoded the phylogenetic relationships, virulence factors, antibiotic resistance genes, and genetic contexts of mcr-1 of 80 isolates, we could not definitively determine how many antibiotic resistance genes coexist with mcr-1 on the same plasmids due to some incomplete assemblies (generated by short reads of Illumina sequencing). SMRT sequencing offers a definitive solution for deciphering individual bacterial isolates by virtue of its long reads, but the significantly higher cost limits its application for large numbers of samples. We therefore subjected 15 isolates to SMRT sequencing to complement the data set generated by Illumina sequencing. This combined-sequencing method revealed that mcr-1 was located on the chromosome of 7 isolates and in the plasmids of 73 isolates. New sequencing technologies (such as nanopore sequencing) that generate long sequence reads may be used in future studies to decipher mcr-1-carrying MDR plasmids (56), which would facilitate our understanding of the cotransfer of mcr-1 with other antimicrobial resistance genes in bacterial plasmids.

**MATERIALS AND METHODS**

**Ethical approval.** Ethical approval of the study was granted by the Second Affiliated Hospital of Zhejiang University.

**Genome sequencing and assembly.** Genomic DNA of all mcr-1-positive isolates was extracted using a Wizard genomic DNA purification kit (Promega, Beijing, China), following the manufacturer’s instructions. Indexed Illumina sequencing libraries were prepared using a TruSeq DNA PCR-free sample preparation kit (Illumina Inc., San Diego, CA) following the standard protocol and were sequenced on an Illumina HiSeq 2500 platform according to the manufacturer’s protocols; the sequencing produced 250-bp paired-end reads (Bionova, Beijing, China). The draft genomes were assembled using the SPAdes algorithm (57) and CLC Genomics workbench 8.5 (CLC Bio, Aarhus, Denmark). To optimize the assembly of plasmids, Illumina reads of each isolate were reassembled by plasmidSPAdes, a program to assemble plasmids from whole-genome sequencing data (58). All the assemblies were further corrected by Pilon (59). The mcr-1-containing contigs for each isolate were extracted from the two assemblies, and the longer one was used for analysis of the genetic context of mcr-1. Although we generated only short mcr-1-containing contigs in Illumina sequencing and 6 isolates that were able to cotransfer other antibiotic resistance genes with mcr-1 were further sequenced by the use of SMRT sequencing (60). Genomic DNA was sheared to 10 to 17 kb using Covaris g-tubes (Covaris) and converted into SMRTbell template libraries. The libraries were subsequently subjected to DNA size selection using a BluePippin instrument (Sage Science) to select the longest DNA fragments (lower size cutoff value of ~5 kb). Sequencing was performed on a PacBio RSII system using P6 polymerase binding and C4 sequencing kits with magnetic bead loading and 120-min acquisition (Sinobiocore, Beijing, China). Genome assemblies were performed using HGAP and Quiver as part of SMRTAnalysis version 2.3 by the use of the HGAP3 protocol and corrected using Pilon. Shotgun sequences of all 80 isolates have been deposited in the NCBI database (see below).

**Molecular epidemiology and analysis of virulence and antibiotic resistance genes.** Assembled genomes from Illumina and Pacbio sequencing were aligned, and a core-genome phylogenetic tree was generated by Parsnp in the Harvest package (61). MLST analysis and examinations of known virulence-associated genes and antibiotic resistance genes were carried out using pipeline SRST2, which takes Illumina reads as the input (62). Reference sequences of virulence genes and antibiotic resistance genes were from databases VFDB (63) and ARG-ANNOT (64), respectively. The tree and the molecular features of each isolate were visualized by the use of the online tool iTOL (65). Classification of strains to phylogenetic groups was performed by using assembled contigs according to a scheme described previously (32, 66).

**Analysis of mcr-1 location.** The plasmid or chromosome location of the mcr-1 gene in 80 E. coli isolates was first determined by S1-nuclease digestion, pulsed-field gel electrophoresis, and probing with the mcr-1 DNA fragment as described previously (9). The location of mcr-1 was further confirmed by sequence analysis. mcr-1-containing contigs generated by Illumina and Pacbio sequencing were examined for Inc types by PlasmidFinder (40). The mcr-1-carrying contigs encoding plasmid replicons were considered to come from plasmids, and all the contigs of the genome of each corresponding isolate were...
mapped to the reference plasmids by BLASTN (67). A closest reference plasmid was selected based on the mapping coverage to scaffold the mcr-1-carrying plasmid from each isolate.

Analysis of the genetic context of mcr-1. The mcr-1-carrying contigs were annotated by the use of the RAST annotation server (68). The insertion sequences were identified by ISfinder (69). For contigs in which mcr-1 was immediately surrounded by IS elements or only the mcr-1-hp core region was observed, an inverse PCR assay using primers F-primer (TATTCTGTGCCGCTGATGTT) and R-primer (TATCCGGCCTG GTGCCTT) (annealing temperature, 55°C) located within the mcr-1 gene was performed to determine the possible existence of a free circular form containing IS-flanked sequence.

Additional material. Supplemental figures can be found at figshare (https://doi.org/10.6084/m9.figshare.6281579.v1). These show the genetic contexts of mcr-1 in the E. coli isolates from one hospital.

Accession number(s). Shotgun sequences of all 80 isolates have been deposited in the NCBI database (BioProject accession no. PRJNA331013 and BioSample accession no. SAMN05437795 to SAMN05437874), while 15 SMRT sequences have been deposited under other accession numbers (BioProject accession no. PRJNA3880845 and BioSample accession no. SAMN06649969 to SAMN06649983).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/mBio.00943-18.

TABLE S1, XLSX file, 0.04 MB.
TABLE S2, XLSX file, 0.05 MB.
TABLE S3, XLSX file, 0.03 MB.
TABLE S4, XLSX file, 0.1 MB.
TABLE S5, XLSX file, 0.03 MB.

ACKNOWLEDGMENTS

This work was supported in part by the National Natural Science Foundation of China (grants 31672604 and 31422055) and the Natural Science Foundation of Zhejiang Province (2013C33G2010505), Higher Education Funding Council and European Commission, United Kingdom.

REFERENCES

17. McGann P, Snesrud E, Maybank R, Corey B, Ong AC, Clifford R, Hinkle M,


