JAC Antimicrob Resist
https://doi.org/10.1093/jacamr/dlad067

JAC-Antimicrobial Resistance

Whole genome sequencing of multidrug resistant Enterobacterales identified in children and their household members within Siem Reap, Cambodia

Shweta R. Singh[1], Cheng Yee Tang[1], Bunsoth Mao[2], Sona Soeng[3], Clare L. Ling[3,5], Jocelyn Qi-Min Teo[4], Saphonn Vonthanak[2], Paul Turner [3,5], Li Yang Hsu [1,6,7] and Rick Twee-Hee Ong[1]*

[1]Saw Swee Hock School of Public Health, Tahir Foundation Building, National University of Singapore, Singapore; [2]University of Health Sciences, Phnom Penh, Cambodia; [3]Cambodia Oxford Medical Research Unit, Angkor Hospital for Children, Siem Reap, Cambodia; [4]Department of Pharmacy, Singapore General Hospital, Singapore; [5]Centre for Tropical Medicine and Global Health, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK; [6]Department of Medicine, Yong Loo Lin School of Medicine, NUHS Tower Block, National University of Singapore, Singapore; [7]Singapore Centre on Environmental Life Sciences Engineering (SCELSE), Nanyang Technological University, Singapore

*Corresponding author: Email: rick.tweehee.ong@nus.edu.sg

Received 21 February 2023; accepted 2 May 2023

Objectives: To explore the association of recent hospitalization and asymptomatic carriage of multidrug-resistant Enterobacterales (MDRE) and determine the prevailing strains and antibiotic resistance genes in Siem Reap, Cambodia using WGS.

Methods: In this cross-sectional study, faecal samples were collected from two arms: a hospital-associated arm consisted of recently hospitalized children (2–14 years), with their family members; and a community-associated arm comprising children in the matching age group and their family members with no recent hospitalization. Forty-two families in each study arm were recruited, with 376 enrolled participants (169 adults and 207 children) and 290 stool specimens collected from participants. The DNA of ESBL- and carbapenemase-producing Enterobacterales cultured from the faecal samples was subject to WGS on the Illumina NovaSeq platform.

Results: Of the 290 stool specimens, 277 *Escherichia coli* isolates and 130 *Klebsiella* spp. were identified on CHROMagar ESBL and KPC plates. The DNA of 276 *E. coli* (one isolate failed quality control test), 89 *Klebsiella pneumoniae*, 40 *Klebsiella quasipneumoniae* and 1 *Klebsiella variicola* was sequenced. CTX-M-15 was the most common ESBL gene found in *E. coli* ($n=104$, 38%), *K. pneumoniae* ($n=50$, 56%) and *K. quasipneumoniae* ($n=16$, 40%). The prevalence of bacterial lineages and ESBL genes was not associated with any specific arm.

Conclusions: Our results demonstrate that MDRE is likely to be endemic within the Siem Reap community. ESBL genes, specifically $bla_{CTX\text{-}M}$, can be found in almost all *E. coli* commensals, indicating that these genes are continuously propagated in the community through various unknown channels at present.

Introduction

The increasing incidence of MDR in bacterial pathogens has emerged as a silent epidemic. MDR pathogens reduce antimicrobial efficacy causing prolonged illness and increasing the burden on strained healthcare systems.[1,2] MDR Enterobacterales (MDRE), including ESBL-producing Enterobacterales (ESBL-E) and carbapenem-resistant Enterobacterales (CRE), are identified as 'critical priority' resistant organisms by the WHO.[3] Initially associated with clinical infections acquired within healthcare institutions, lately ESBL-E have increasingly been found in community-acquired infections.[4]

Klebsiella pneumoniae and *Escherichia coli* are two important commensal organisms found in the human gut and are also common causes of extra-intestinal infections. Carriage of ESBL-E has been identified as a risk factor for clinical infection.[5,6] Recent reviews on faecal ESBL-E carriage have shown an increasing trend in the ESBL-E carriage rates in healthy individuals worldwide and specifically in the Southeast Asian region.[7,8] To control MDRE spread, it is essential to identify the reservoirs and dissemination paths of resistant bacteria and antibiotic resistance genes (ARGs) that have contributed to this public health crisis.

In Cambodia, several studies have been conducted to investigate the incidence of MDRE; a study conducted on clinical

specimens reported increasing proportions of ESBL-E infections, from 24% in 2012% to 38% in 2015.[9] Another study in the same hospital's outpatient department showed that 55% of children were colonized by extended-spectrum cephalosporin-resistant Enterobacterales (*E. coli* and *K. pneumoniae*).[10] Risk factors associated with colonization included recent hospitalization and presence of intestinal parasites. However, in a recent cross-sectional study conducted in 2019, there was no difference in ESBL-E colonization prevalence in recently hospitalized children and their families compared with children and individuals with no recent history of hospitalization. The prevalence of ESBL-E colonization (92% ESBL *E. coli* and 44% *K. pneumoniae*) reported in the study was higher compared with all previous studies in Cambodia.[11] The aim of this study was therefore to determine the population structures of the bacterial strains and ARGs present in *E. coli* and *K. pneumoniae*, particularly of ESBL genes from the 2019 study, which could identify potential reservoirs of ARGs.

Methods

Patients and setting

Faecal samples were collected from community participants of Siem Reap, Cambodia in a cross-sectional study from August to November 2019. The study included two arms: a hospital-associated household (where a child aged 2–14 years was hospitalized in the past 14–28 days for a minimum 48 h); and a community-associated household, where a child in a matching age group (2–14 years) and other family members were not hospitalized in the past 12 months. The children in the hospital-associated arm were identified and recruited from Angkor Hospital for Children (AHC) in Siem Reap and approached between 14 and 28 days after discharge. Data collectors interviewed consenting residents who had lived in the household for at least 1 month before the recruitment day, using a tablet-based questionnaire on the Qualtrics platform. We obtained parental consent for participants under 18 years old and separate assent for children aged 7 years and above. We made repeated visits to each household to recruit additional household members if needed, and there were no exclusion criteria for other residents in the households.

Households were approached in the same street or village as the hospital-associated households that had a child of similar age (±2 years) to the index child. On the same day, households where the age-matched child and any other family member had not been hospitalized in the past year were eligible for inclusion in the study as community-associated households. We recruited the first eligible community household that agreed to participate in the study.

Microbiological methods

The 290 faecal samples were stored at −80°C in tryptone soya broth/10% glycerol medium immediately on receipt at the laboratory and thawed for interval processing in batches. The samples were cultured to detect ESBL-producing *E. coli* and *K. pneumoniae* using chromogenic media (ESBL, KPC agar; CHROMagar, Paris, France; prepared in-house). Suspect colonies (one distinct colony per sample) were confirmed with MALDI-TOF (VITEK MS, database V3.2; bioMérieux, Marcy L' Etoile, France). Antimicrobial susceptibilities were examined by disc diffusion, and zone diameters were interpreted using the CLSI guidelines, 2019 version.[12] ESBL and carbapenemase production were confirmed using double-disc diffusion and the modified carbapenem inactivation method (mCIM), respectively. Non-susceptibility was defined as isolates resistant or intermediate susceptible to antimicrobials. The subcultured isolates (ESBL and carbapenemase-positive *E. coli* and *K. pneumoniae*) were stored frozen at −80°C in skim milk/tryptone/glucose/glycerol medium.

WGS

DNA was extracted from subcultured isolates using the Promega Wizard Genomic DNA purification kit, following the manufacturer's instructions. The DNA samples were then tested for purity and integrity using agarose gel electrophoresis followed by Qubit 3.0 fluorometer quantitation. Samples with a DNA yield greater than 400 ng were library prepped using the NEBNext® Ultra™ II DNA Library Prep Kit. Post further quality checks and purification, pooled DNA libraries were sequenced on the Illumina® NovaSeq 6000 sequencing platform to yield 150 bp paired-end reads.

Bioinformatics analysis

The raw sequence reads were trimmed to remove adaptors and low-quality bases using Trimmomatic (v0.32),[13] and were subsequently de novo assembled using SPAdes (v3.11.1).[14] Species identification was performed using Mash Screen (v2.3) by comparing the raw sequencing reads with a set of species database curated by the Kleborate (downloaded on 23 June 2021).[15] All samples were subjected to *in silico* MLST against the PubMLST database (https://pubmlst.org) (downloaded on 31 August 2021) using a custom Python script integrating BLASTn. The Achtman MLST scheme was used for *E. coli*, whereas the Pasteur MLST scheme was used for *K. pneumoniae*, *K. quasipneumoniae* and *K. variicola*. The *E. coli* isolates were classified in phylogenetic groups based on the presence/absence of genes *chuA*–, TspE4.C2– (Group A); *chuA*–, *yjaA*–, TspE4.C2+ (Group B1); *chuA*+, *yjaA*+ (Group B2); and *chuA*+, *yjaA*– (Group D).[16] The presence of ARGs in the assembled nucleotide sequences was identified using AMRFinderPlus (v3.10.1) with its own database (version 2021-03-01.1) using the gene sequence coverage and identity threshold set as >80% and >90%, respectively.[17] All genome assemblies were subjected to additional genotyping using Kleborate (v2.0.4).[18,19] Plasmid incompatibility groups were identified using PlasmidFinder (v2.0.1) with its own database (version 2021-07-12).[20] The genetic context of specific bla_{CTX-M} and bla_{NDM-5} genes was examined by extracting the contigs containing these genes and annotating the contigs using Prokka (v1.14.6).[21] ISs in the contigs were verified using IS finder.[22] The contigs that contained the same bla_{CTX-M} allele variant were clustered using CD-HIT (v4.8.1)[23] and verified using BLASTn. Schematics of the gene features were visualized using R (v4.1.0) ggplot2[24] and gggenes packages.[25]

Comparison of genes between the two arms (hospital and community) for the presence or absence of certain genes was tested using chi-square test. Fisher's exact test was used when the sample size of prevalent genes was below 30.

For phylogenetic analysis, trimmed reads were aligned to an appropriate reference genome: AE014075.1 *E. coli* CGT073 strain or CP000647.1 *K. pneumoniae* subsp. *pneumoniae* MGH 78578 strain. The core genome alignments were generated using Snippy (v4.6.0).[21] Recombination sites were removed using Gubbins (allowing 28% missing data threshold for *E. coli* and 25% missing data threshold for *K. pneumoniae*).[26] Pairwise SNP distance was calculated from the recombination-free core genome alignment using snp-dists.[27] Maximum likelihood trees were constructed from the recombination-free core genome alignment with IQ-TREE (v2.0.3)[28] and visualized using R ggplot2 and ggtree packages.[29]

Results

Stool samples were collected from 290 participants in 84 families (42 family units in each arm). A total of 277 *E. coli* were identified on the CHROMagar plates of which 269 (92.8%, 269/290) were ESBL positive and 5 isolates were ESBL negative but third-generation cephalosporin-resistant. Four *E. coli* isolates were positive for carbapenemase, of which one was also positive for ESBL. *K. pneumoniae* was isolated from 131 ESBL and KPC CHROMagar plates, of which 128 (44.1%, 128/290) were confirmed as ESBL-positive and 3 as mCIM-positive. Two of the three mCIM-positive *K. pneumoniae* isolates came from the same

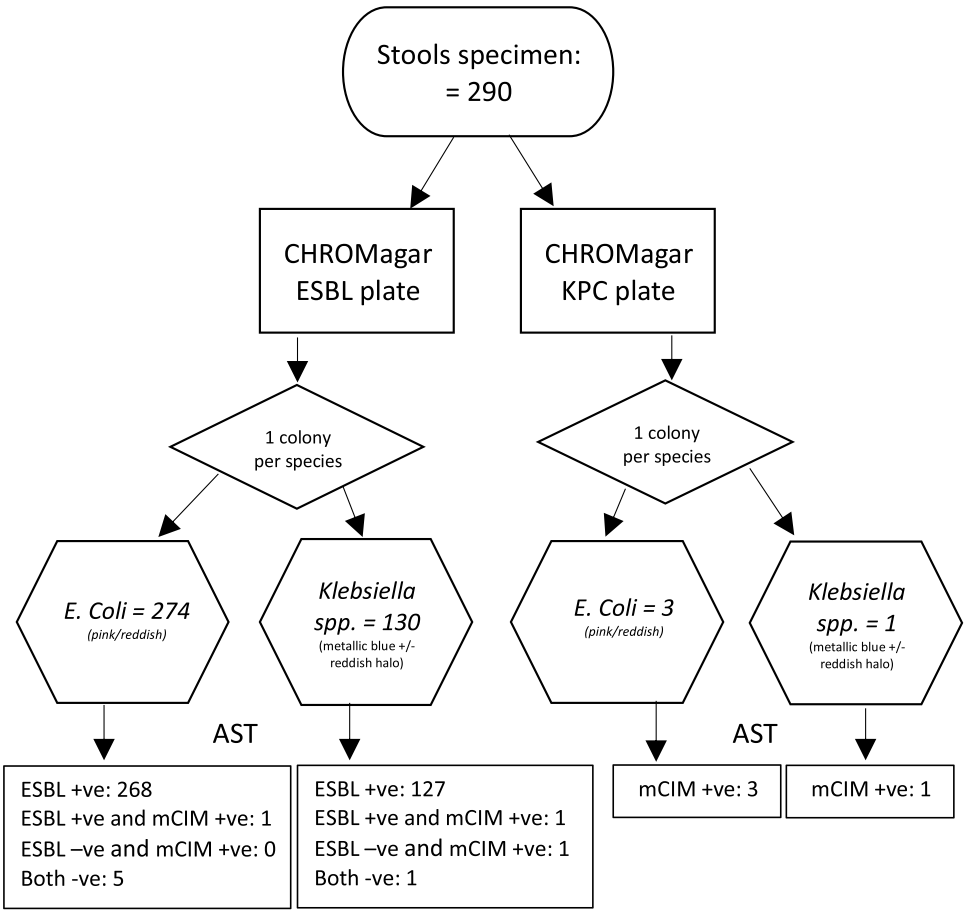


Figure 1. Phenotype detection of ESBL *E. coli* and *Klebsiella* spp. isolates. mCIM, modified carbapenem inactivation method.

individual (cultured from different CHROMagar plates) and hence only one isolate was used for further processing. All of the 407 (277 *E. coli* and 130 *K. pneumoniae*) identified isolates had their DNA extracted and submitted for WGS (Figure 1). Of these, the DNA of one *E. coli* isolate failed to meet the DNA quality threshold and therefore sequence data were available for 406 isolates. The fastq reads for the 406 isolates have been uploaded to NCBI, and the project assession number is PRJNA885285.

Geontypic species identification from the sequence data confirmed all 276 phenotypic *E. coli* samples as *E. coli*/Shigella (273) and *Escherichia* cryptic clade-1 (3). For the 130 phenotypic *K. pneumoniae* isolates, Kleborate identified 89 to be *K. pneumoniae*, 40 as *K. quasipneumoniae* and 1 as *K. variicola*. The prevalent *E. coli* STs were ST10 ($n=17$), ST131 ($n=15$), ST1193 ($n=14$) and ST38 ($n=13$), as shown in Table S1 (available as Supplementary data at *JAC-AMR* Online). The major STs for *K. pneumoniae* isolates were ST101 ($n=9$), ST37 ($n=4$), ST1741 ($n=4$) and ST967 ($n=4$) whereas for *K. quasipneumoniae* the major STs were ST1308 ($n=8$), ST841 ($n=4$) and ST1584 ($n=4$), as shown in Table S2. There were no differences in proportions of the STs for any of the ESBL-E (*E. coli*, *K. pneumoniae* and *K. quasipneumoniae*) between the hospital-associated households and community-associated households.

The gene $bla_{\text{CTX-M-15}}$ was the most commonly found ESBL gene in *E. coli* ($n=104$, 38%), *K. pneumoniae* ($n=50$, 56%) and *K. quasipneumoniae* ($n=16$, 40%). The second most common ESBL gene in *E. coli* was $bla_{\text{CTX-M-55}}$ ($n=90$/32%), followed by $bla_{\text{CTX-M-27}}$ ($n=48$/17%). The $bla_{\text{CMY-42}}$ gene was found in all five *E. coli* ESBL-negative third-generation cephalosporin-resistant isolates, whereas the bla_{SHV} gene and its variants were present in one ESBL-negative and four phenotypically ESBL-positive *K. pneumoniae*. The distribution of the ESBL genes across the three species is summarized in Tables 1 and 2. CTX-M-27 was most commonly found in *E. coli* with ST131 and ST1193, although the association was not statistically significant (Table S3). *K. pneumoniae* and *K. quasipneumoniae* distribution of major ST and ESBL genes is given in Table S4.

The $bla_{\text{NDM-5}}$ gene was present in all of the six carbapenemase-producing (CP) Enterobacterales (four *E. coli* and two *K. pneumoniae*). The STs for CP *E. coli* isolates were ST167, ST410 ($n=2$) and ST648. One *E. coli* ST131 isolate was identified to harbour $bla_{\text{OXA-48}}$ from the sequence data, which is not surprising given that CHROMagar KPC plates are not 100% sensitive or specific.[30] The STs of the two CP *K. pneumoniae* isolates were ST45 and ST11.

Likely transmissions within the community

To identify potential transmissions, we assessed the core phylogenies of the *E. coli* and *K. pneumoniae* isolates as illustrated in

Table 1. Distribution of ESBL *E. coli* ($n=276$) phenotypes and major genes

Genes	Hospital arm	Community arm	Total	*P* value[a]
CTX-M-15	57	47	104	0.37
CTX-M-27	33	17	50	0.32
CTX-M-14	13	8	21	0.93
CTX-M-55	52	38	90	0.93
CTX-M-9	—	—	—	
CTX-M-3	0	1	1	
CTX-M-24	2	1	3	
CTX-M-65	1	2	3	
CTX-M-63	—	—	—	
bla_{CMY-2}-positive[b]	6	0	6	
bla_{SHV}-positive				
SHV-2/2-like	—	—	—	
SHV-12	—	—	—	
SHV-26[c] + 238S	—	—	—	
Carbapenemase-positive				
bla_{NDM-5}-positive	4	0	4	
bla_{OXA-48}-positive	1	0	1	

[a]Chi-square/Fisher's exact test.
[b]Three bla_{CMY-2}-positive *E. coli* also possessed CTM-M-15 and CTX-M-27 genes.
[c]Three *E. coli* isolates possessed more than two types of ESBL acquired genes.

Table 2. ESBL-*K. pneumonia* and ESBL-*K. quasipneumoniae* phenotypes and major genes

	K. pneumoniae ($n=90$)				*K. quasipneumoniae* ($n=40$)			
Genes	H	C	Total	*P* value[b]	H	C	Total	*P* value[a]
CTX-M-15	29	21	50	0.86	11	5	16	0.41
CTX-M-27	3	2	5		7	3	10	0.44
CTX-M-14	3	3	6		2	3	5*	
CTX-M-55	5	2	7		2	1	3	
CTX-M-9	1	4	5		1	1	2	
CTX-M-3[b]	2	2	4		0	1	1	
CTX-M-24	0	2	2		—	—	—	
CTX-M-65	0	0	0		—	—	—	
CTX-M-63	1	0	1		—	—	—	
SHV-2/2-like	1	2	3		0	2	2	
SHV-12	—	—	—		0	2	2	
SHV-26* + 238S	1	0	1		—	—	—	
Carbapenemase-positive								
bla_{NDM-5}-positive	2	0	2		—	—	—	
bla_{OXA-48}-positive	—	—	—		—	—	—	

H, Hospital arm; C, Community arm.
[a]Chi-square/Fisher's exact test.
[b]The one *K. variicola* isolate had the CTX-M-3 ESBL gene.

Figures 2 and 3. Using a threshold of 10 SNPs, we identified: (i) 18 instances of potential transmission for *E. coli* of which 10 instances were from individuals in nine families (Table S5); and (ii) 11 instances of potential transmission for *K. pneumoniae* in seven families (Table S6). In the 18 instances of potential transmission for *E. coli*, the majority consisted of the combination of ST131/1193 with CTX-M-27 (three instances) found both within and across different households. This was followed by ST10 and ST215 with the ESBL gene CTX-M-15 (Table 3). With respect to *K. pneumoniae*, ST101 with CTX-M-15 was found in two separate possible transmission instances in individuals belonging to different households.

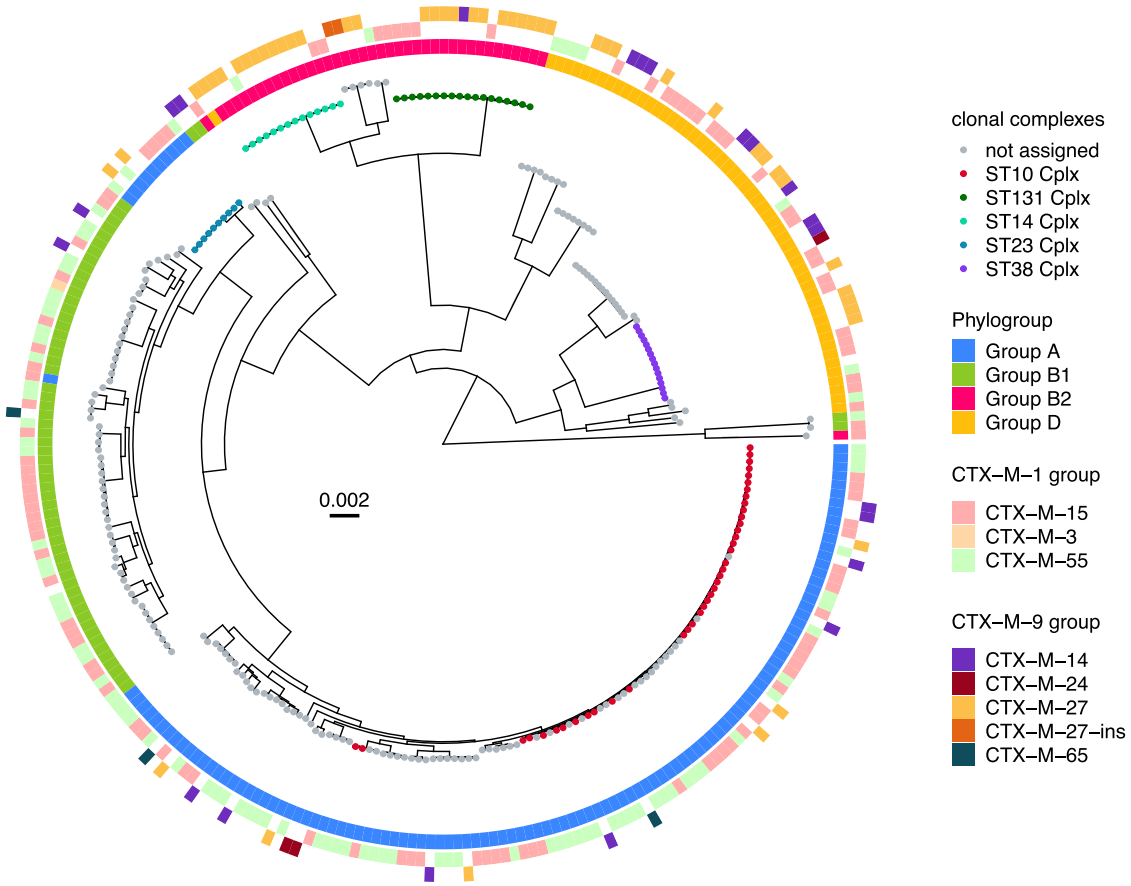


Figure 2. Recombination-filtered core SNP maximum likelihood phylogenetic tree for 277 *E. coli* isolates. The tree was midpoint-rooted. The tips were coloured by the ST clonal complexes. Inner circle shows the phylogroup of the isolates; middle circle shows the presence of CTX-M-1 group genes identified in the isolates; outer circle shows the presence of CTX-M-9 group genes identified in the isolates. Scale bar indicates the number of nucleotide substitutions per site.

Presence of other ARGs besides ESBL

WGS also enabled the identification of genes that are associated with resistance to several antibiotics including ciprofloxacin, co-trimoxazole, chloramphenicol and gentamicin (Table 3). Ciprofloxacin non-susceptible (NS) *E. coli* and *K. pneumoniae* isolates most commonly possessed the genes *qnrS1/S1**. *E. coli* isolates non-susceptible to co-trimoxazole were found to harbour the sulfamethoxazole resistance gene *sul2/2*/2^* ($n=52$, 27.22%), *sul1/1** ($n=40$, 21%) and their combination *sul1*;*sul2* ($n=60$, 31.4%). The gene *sul2* was also most prevalent in NS *K. pneumoniae* ($n=46$, 76.7%) followed by *Sul1* ($n=35$, 58.3%). However, a number of NS isolates (nine *E. coli* and three *K. pneumoniae*) were not found to harbour any sul- or tmt (trimethoprim)-specific genes, suggesting that there might be other mechanisms for resistance. Colistin-resistant genes *mcr3.1* ($n=4$) and *mcr1.1* ($n=2$) were found in five *E. coli* isolates (one isolate possessed both colistin resistance genes). The phenotypic susceptibilities to major classes of antibiotics—ciprofloxacin, co-trimoxazole, chloramphenicol and gentamicin—and their corresponding genes (identified using Kleborate) are presented in Tables S7-S9.

Genetic context of $bla_{CTX\text{-}M\text{-}27}$

The *E. coli* isolates possessing ESBL $bla_{CTX\text{-}M\text{-}27}$ genes were further analysed for their genetic context. Of these 50 isolates, 18 (36%) were likely to be on plasmids and 4 (8%) were likely of chromosomal origin, whereas locales could not be adequately determined for the remaining isolates (Figure S1). The CTX-M-27-containing contigs are unfortunately not typeable, whereas the remaining contigs predicted to be from plasmids (lengths ranging from 1470 to 6590 bp) had similar sequence identities (>90% similarity and 100% query coverage) to several plasmid sequences such as CP054458.1 and CP088684.1. All of the E.coli $bla_{CTX\text{-}M\text{-}27}$ genes were found to harbour IS*Ecp1* upstream, with lengths ranging from 387 to 540 bp in 30 isolates and truncated in the remaining 10 isolates. IS*903B* was consistently found downstream, with lengths ranging from 388 to1057 bp in 37 isolates (only three did not have the IS*903B*). The *K. pneumoniae* and *K. quasipneumoniae* isolates with $bla_{CTX\text{-}M\text{-}27}$ also showed similar genetic structure with IS*Ecp1* upstream and IS*903B* downstream, with lengths of 1656 and 598 bp or unknown respectively (Figure S2). For *K. pneumoniae*, the plasmid

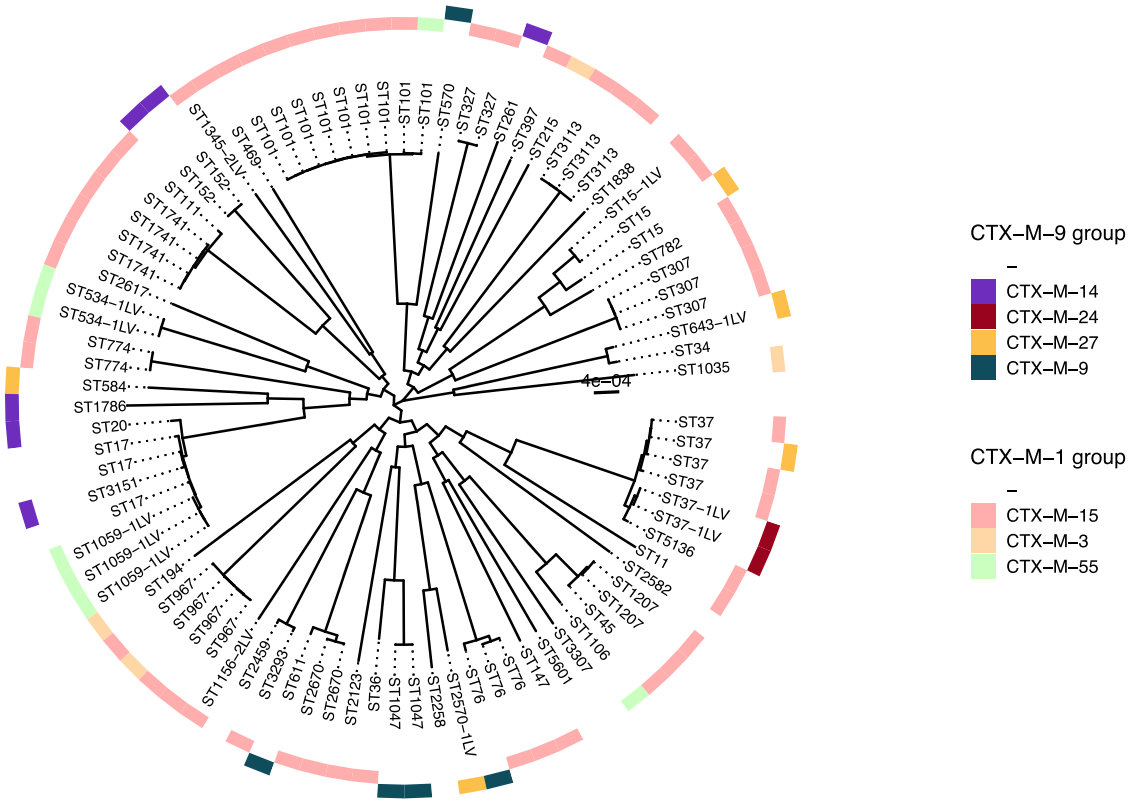


Figure 3. SNP phylogenetic circular tree for 131 *K. pneumoniae* isolates. The tree was midpoint-rooted. The tips were coloured by the ST clonal complexes. Inner circle shows the phylogroup of the isolates; middle circle shows the presence of CTX-M-1 group genes identified in the isolates; outer circle shows the presence of CTX-M-9 group genes identified in the isolates. Scale bar indicates the number of nucleotide substitutions per site.

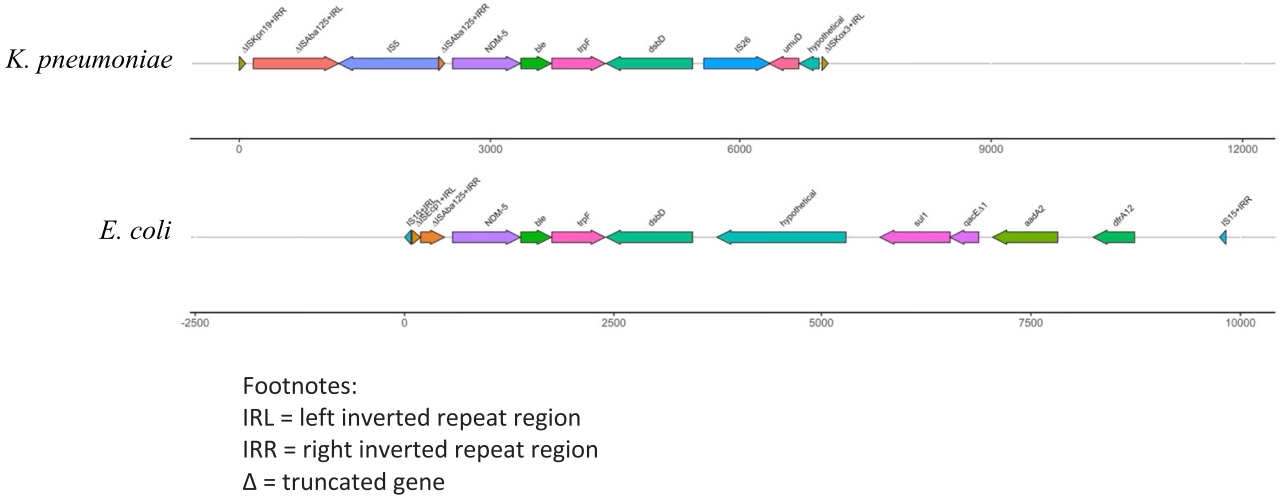
Table 3. Distribution of prevalent AMR genes present in non-susceptible *E. coli* and *K. pneumoniae*

	E. coli ($n=276$)			*K. pneumoniae* ($n=89$)		
Antibiotics	Total (%)	Common genes	Total (%)	Total (%)	Common genes	Total (%)
Colistin	5 (1.8)	mcr3.1	4 (80)	—	—	
		mcr1.1	2 (40)			
Ciprofloxacin	252 (91.3)	qnrS1/S1*	132 (52.4)	83 (93.3)	qnrS1/S1*	58 (69.9)
Co-trimoxazole/sulfamethoxazole/trimethoprim	191 (69.2)	sul2 and/or sul1	152 (79.6)	60 (67.4)	sul2/S*	46 (76.7)
		dfrA17	74 (38.7)		dfrA14.v2*	25 (41.7)
Chloramphenicol	87 (31.5)	floR.v1/v1*	22 (25.3)	23 (25.8)	floR.v1	11 (47.8)
		catII.2*	17 (19.5)			
		catA1*/A1^	16 (18.4)			
Gentamicin	82 (29.7)	aac(3)-iid^	45 (54.9)	19 (21.3)	aac(3)-iid^	11 (57.9)

replicon IncFIA(HI1) was identified to be present within a CTX-M-27-containing contig only.

Genetic context of bla_{NDM-5} in isolates from same individual

The genetic structure of bla_{NDM-5} found in *E. coli* and *K. pneumoniae* from the same individual (an infant with no medical or hospitalization history in the past 1 year) was also analysed. The contig length harbouring bla_{NDM-5} in the *E. coli* isolate differed (9827 bp) substantially from the *K. pneumoniae* isolate (7054 bp), where only 49% of the *K. pneumoniae* and 35% of the *E. coli* bla_{NDM-5} contigs matched, indicating that the genetic environment was not similar between *E. coli* and *K. pneumoniae* (Figure 4). The assembled contigs were, however, identical within the same species isolated from different individuals.



Footnotes:
IRL = left inverted repeat region
IRR = right inverted repeat region
Δ = truncated gene

Figure 4. Schematic of the genetic structure of the contigs harbouring NDM-5 in *K. pneumoniae* and *E. coli* isolated from an individual.

Discussion

The WGS and subsequent analysis showed that there was no significant difference in the ESBL genes and MDRE strains depending on the arm, and thus recent hospitalization has no association with a specific ESBL gene or strain. The significant prevalence and substantial genetic diversity of the ESBL genes within Siem Reap highlight that ESBL-producing *E. coli* and *K. pneumoniae* are likely to be endemic in the community. The results from this study also indicate that there are no major differences in the stool specimen-cultured *E. coli* and *K. pneumoniae* genotypes in families of recently hospitalized children compared with families with children that had no recent hospitalization. Our study contrasts with findings from another study from Spain where the index patient and their household members had higher prevalence of ESBL-E compared with non-hospitalized community individuals.[31] The bacterial strains and ESBL genes were also very diverse in our sample, with several instances of possible transmission identified. The possible transmissions were, however, distinct across the clusters harbouring different sequence types and ESBL genotypes. There were several transmission clusters with known *E. coli* strains such as ST131 and ST1193 that had been identified as 'high-risk clones', spreading across all regions in the world.[32–34] In two separate transmission clusters, *K. pneumoniae* ST101 was identified, hinting at a common environmental link of possible transmission. Previous studies have linked a *K. pneumoniae* ST101 clone with not only harbouring carbapenemase resistance but also persisting in the environment through hospital wastewater effluent.[35,36] The high pervasiveness and isolation from unrelated individuals warrants further inspection of environmental sources that could be disseminating the 'high-risk clones' in the community.

The predominant ESBL gene family in our isolates was $bla_{\text{CTX-M}}$, with major allelic variants being $bla_{\text{CTX-M-15}}$, $bla_{\text{CTX-M-55}}$ and $bla_{\text{CTX-M-27}}$, which are also commonly found elsewhere in the Greater Mekong region.[8,9] We also identified carbapenem resistance in four *E. coli* and two *K. pneumoniae* isolates, all of which possessed the $bla_{\text{NDM-5}}$ gene. In this study, all but one individual with NDM-5-positive isolates had no history of a recent hospitalization. We also identified a carbapenem (meropenem)-susceptible *E. coli* isolate that possessed the $bla_{\text{OXA-48}}$ gene on sequencing. Although CHROMagar KPC plates are not 100% sensitive, especially to OXA-48-like carbapenemase, it is surprising that the *E. coli* strain possessing OXA-48-like carbapenemase was susceptible to meropenem, implying that the OXA-48 gene is not expressed phenotypically.[30] The species-specific $bla_{\text{NDM-5}}$ contigs were identical thus hinting at potential transmission from a common source, particularly because all the positive isolates came from individuals in different households. However, we could not identify a common source of transmission or a link between the individuals as they belonged to different villages and had no history of recent hospitalization. Given the limited number of positive samples, we were not able to identify any independent risk factors associated with colonization by a $bla_{\text{NDM-5}}$-containing strain. With respect to *K. pneumoniae*, we did not identify high-risk clones of *K. pneumoniae* such as ST15[37] and ST147[38] in our sample population. However, it was alarming to identify high-risk clones of ST101[36] and ST11[39] in the community, especially ST101, which was also found in possible transmission events from two separate instances. ST101 is especially known as a 'superbug' that carries both virulence and a combination of carbapenem and colistin resistance genes and poses a grave threat if it spreads widely in the community.[35,36]

One major limitation of this study was that we could only generate short-read genome sequences. The use of long-read sequencing would have enabled determination of the genetic environment harbouring the NDM-5 gene within each bacterial species to identify if it was carried on a species-specific plasmid or within the chromosome. Another limitation was the cross-sectional design, which limited our ability to track the longitudinal differences in transmission dynamics post hospitalization within households. In addition, further studies from the environment and food sources are required to identify the reservoirs in the community and mechanisms of dissemination of antimicrobial

resistant bacteria to the wider population. Recent studies have shown that MDR healthcare-associated bacteria share mobile genetic elements (MGEs) across large phylogenetic distances.[40–42] Identifying the dynamics of MGE transmission in community settings can explain important epidemiological links that remain unsolved.[43,44]

Conclusion

Our study shows that the prevalence of ESBL *E. coli* and *K. pneumoniae*/*K. quasipneumoniae* in Cambodia has not only increased considerably compared with previous reports but is also diverse genotypically. ESBL genes, specifically $bla_{\text{CTX-M}}$, are now found in almost all *E. coli* commensals, indicating that they are continuously propagated in the community through various currently unknown channels. Another startling finding was the presence of $bla_{\text{NDM-5}}$ in both *E. coli* and *K. pneumoniae* from healthy individuals. If carbapenem resistance prevalence increases at a similar rate as was seen for ESBL, then untreatable Enterobacterales will be commonplace in Cambodia in the near future. Effective community-level interventions are required urgently.

Acknowledgements

We thank the team of University Health Sciences, COMRU team at the Angkor Hospital for Children and NUS team for their efforts towards the completion of this project.

Funding

This work was funded by the Ministry of Education, Singapore under the Academic Research Fund Tier 1 (FY2019) and supported by the Singapore Ministry of Health's National Medical Research Council under its Centre Grant Programme (MOH-001010-00). This research was funded in whole, or in part, by the Wellcome Trust (grant number 220211). For the purpose of open access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Transparency declarations

S.R.S. and R.T.-H.O., the lead author and corresponding authors, respectively, and all other authors affirm that this manuscript is an honest, accurate and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained. The authors have no conflicts of interest to declare.

Ethical approval

The study protocol was approved by the National Ethics Committee for Health Research-Cambodia (148 NECHR), Oxford Tropical Research Ethics Committee (OxTREC; 551-18), and National University of Singapore—Institutional Review Board (NUS-IRB; H-18-069).

Supplementary data

Figures S1 and S2 and Tables S1 to S9 are available as Supplementary data at *JAC-AMR* Online.

References

1 Naylor NR, Atun R, Zhu N *et al.* Estimating the burden of antimicrobial resistance: a systematic literature review. *Antimicrob Resist Infect Control* 2018; **7**: 58. https://doi.org/10.1186/s13756-018-0336-y

2 Teerawattanapong N, Panich P, Kulpokin D *et al.* A systematic review of the burden of multidrug-resistant healthcare-associated infections among intensive care unit patients in Southeast Asia: the rise of multidrug-resistant *Acinetobacter baumannii*. *Infect Control Hosp Epidemiol* 2018; **39**: 525–33. https://doi.org/10.1017/ice.2018.58

3 WHO. WHO publishes list of bacteria for which new antibiotics are urgently needed. 2017. https://www.who.int/news/item/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed

4 Holt KE, Wertheim H, Zadoks RN *et al.* Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in *Klebsiella pneumoniae*, an urgent threat to public health. *Proc Natl Acad Sci U S A* 2015; **112**: E3574–E81. https://doi.org/10.1073/pnas.1501049112

5 Luvsansharav U-O, Hirai I, Nakata A *et al.* Prevalence of and risk factors associated with faecal carriage of CTX-M β-lactamase-producing Enterobacteriaceae in rural Thai communities. *J Antimicrob Chemother* 2012; **67**: 1769–74. https://doi.org/10.1093/jac/dks118

6 Zhang H, Zhou Y, Guo S *et al.* High prevalence and risk factors of fecal carriage of CTX-M type extended-spectrum beta-lactamase-producing Enterobacteriaceae from healthy rural residents of Taian, China. *Front Microbiol* 2015; **6**: 239.

7 Karanika S, Karantanos T, Arvanitis M *et al.* Fecal colonization with extended-spectrum beta-lactamase-producing Enterobacteriaceae and risk factors among healthy individuals: a systematic review and metaanalysis. *Clin Infect Dis* 2016; **63**: 310–18. https://doi.org/10.1093/cid/ciw283

8 Singh SR, Teo AKJ, Prem K *et al.* Epidemiology of extended-spectrum beta-lactamase and carbapenemase-producing Enterobacterales in the greater Mekong subregion: a systematic-review and meta-analysis of risk factors associated with extended-spectrum beta-lactamase and carbapenemase isolation. *Front Microbiol* 2021; **12**: 695027.

9 Caron Y, Chheang R, Puthea N *et al.* Beta-lactam resistance among Enterobacteriaceae in Cambodia: the four-year itch. *Int J Infect Dis* 2018; **66**: 74–9. https://doi.org/10.1016/j.ijid.2017.10.025

10 van Aartsen JJ, Moore CE, Parry CM *et al.* Epidemiology of paediatric gastrointestinal colonisation by extended spectrum cephalosporin-resistant *Escherichia coli* and *Klebsiella pneumoniae* isolates in north-west Cambodia. *BMC Microbiol* 2019; **19**: 59. https://doi.org/10.1186/s12866-019-1431-9

11 Singh SR, Mao B, Evdokimov K *et al.* Prevalence of MDR organism (MDRO) carriage in children and their household members in Siem Reap province, Cambodia. *JAC-Antimicrobial Resist* 2020; **2**: dlaa097. https://doi.org/10.1093/jacamr/dlaa097

12 CLSI. *Performance Standards for Antimicrobial Susceptibility Testing—Twenty-Ninth Edition: M100-S29*. 2019.

13 Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014; **30**: 2114–20. https://doi.org/10.1093/bioinformatics/btu170

14 Prjibelski A, Antipov D, Meleshko D *et al.* Using SPAdes de novo assembler. *Curr Protoc Bioinformatics* 2020; **70**: e102. https://doi.org/10.1002/cpbi.102

15 Ondov BD, Starrett GJ, Sappington A *et al.* Mash screen: high-throughput sequence containment estimation for genome discovery. *Genome Biol* 2019; **20**: 232. https://doi.org/10.1186/s13059-019-1841-x

16 Clermont O, Bonacorsi S, Bingen E. Rapid and simple determination of the *Escherichia coli* phylogenetic group. *Appl Environ Microbiol* 2000; **66**: 4555–8. https://doi.org/10.1128/AEM.66.10.4555-4558.2000

17 Feldgarden M, Brover V, Gonzalez-Escalona N *et al*. AMRFinderplus and the reference gene catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. *Sci Rep* 2021; **11**: 12728. https://doi.org/10.1038/s41598-021-91456-0

18 Lam M, Wick RR, Watts SC *et al*. A genomic surveillance framework and genotyping tool for *Klebsiella pneumoniae* and its related species complex. *Nat Commun* 2021; **12**: 1–16. https://doi.org/10.1038/s41467-021-24448-3

19 Wyres KL, Wick RR, Gorrie C *et al*. Identification of *Klebsiella* capsule synthesis loci from whole genome data. *Microbial Genomics* 2016; **2**: e000102. https://doi.org/10.1099/mgen.0.000102

20 Carattoli A, Zankari E, García-Fernández A *et al*. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother* 2014; **58**: 3895–903. https://doi.org/10.1128/AAC.02412-14

21 Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014; **30**: 2068–9. https://doi.org/10.1093/bioinformatics/btu153

22 Siguier P, Perochon J, Lestrade L *et al*. ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res* 2006; **34** Suppl 1: D32–D6. https://doi.org/10.1093/nar/gkj014

23 Fu L, Niu B, Zhu Z *et al*. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 2012; **28**: 3150–2. https://doi.org/10.1093/bioinformatics/bts565

24 Team RC. *R: a language and environment for statistical computing. R Foundation for Statistical Computing,* 2016. http://www.R-project.org/

25 Github. *gggenes* 2022. https://github.com/wilkox/gggenes.

26 Croucher NJ, Page AJ, Connor TR *et al*. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using gubbins. *Nucleic Acids Res* 2015; **43**: e15. https://doi.org/10.1093/nar/gku1196

27 Github. tseemann/snp-dists. https://github.com/tseemann/snp-dists/find/master

28 Nguyen L-T, Schmidt HA, von Haeseler A *et al*. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 2014; **32**: 268–74. https://doi.org/10.1093/molbev/msu300

29 Yu G, Smith DK, Zhu H *et al*. Ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol Evol* 2017; **8**: 28–36. https://doi.org/10.1111/2041-210X.12628

30 Moubareck CA, Hammoudi Halat D, Sartawi M *et al*. Assessment of the performance of CHROMagar KPC and Xpert carba-R assay for the detection of carbapenem-resistant bacteria in rectal swabs: first comparative study from Abu Dhabi, United Arab Emirates. *J Glob Antimicrob Resist* 2020; **20**: 147–52. https://doi.org/10.1016/j.jgar.2019.07.021

31 Torres E, López-Cerero L, Morales I *et al*. Prevalence and transmission dynamics of *Escherichia coli* ST131 among contacts of infected community and hospitalized patients. *Clin Microbiol Infect* 2018; **24**: 618–23. https://doi.org/10.1016/j.cmi.2017.09.007

32 Chen SL, Ding Y, Apisarnthanarak A *et al*. The higher prevalence of extended spectrum beta-lactamases among *Escherichia coli* ST131 in Southeast Asia is driven by expansion of a single, locally prevalent subclone. *Sci Rep* 2019; **9**: 13245. https://doi.org/10.1038/s41598-019-49467-5

33 Pitout JDD, Laupland KB. Extended-spectrum β-lactamase-producing Enterobacteriaceae: an emerging public-health concern. *Lancet Infect Dis* 2008; **8**: 159–66. https://doi.org/10.1016/S1473-3099(08)70041-0

34 Taati Moghadam M, Mirzaei M, Fazel Tehrani Moghaddam M *et al*. The challenge of global emergence of novel colistin-resistant *Escherichia coli* ST131. *Microbial Drug Resistance* 2021; **27**: 1513–24. https://doi.org/10.1089/mdr.2020.0505

35 Popa LI, Gheorghe I, Barbu IC *et al*. Multidrug resistant *Klebsiella pneumoniae* ST101 clone survival chain from inpatients to hospital effluent after chlorine treatment. *Front Microbiol* 2020; **11**: 610296. https://doi.org/10.3389/fmicb.2020.610296

36 Roe CC, Vazquez AJ, Esposito EP *et al*. Diversity, virulence, and antimicrobial resistance in isolates from the newly emerging *Klebsiella pneumoniae* ST101 lineage. *Front Microbiol* 2019; **10**: 542. https://doi.org/10.3389/fmicb.2019.00542

37 Zhou K, Lokate M, Deurenberg RH *et al*. Use of whole-genome sequencing to trace, control and characterize the regional expansion of extended-spectrum β-lactamase producing ST15 *Klebsiella pneumoniae*. *Sci Rep* 2016; **6**: 20840. https://doi.org/10.1038/srep20840

38 Pan Y-J, Lin T-L, Lin Y-T *et al*. Identification of capsular types in carbapenem-resistant *Klebsiella pneumoniae* strains by *wzc* sequencing and implications for capsule depolymerase treatment. *Antimicrob Agents Chemother* 2015; **59**: 1038–47. https://doi.org/10.1128/AAC.03560-14

39 Yan JJ, Wang MC, Zheng PX *et al*. Associations of the major international high-risk resistant clones and virulent clones with specific ompK36 allele groups in *Klebsiella pneumoniae* in Taiwan. *New Microbes New Infect* 2015; **5**: 1–4. https://doi.org/10.1016/j.nmni.2015.01.002

40 Cerqueira GC, Earl AM, Ernst CM *et al*. Multi-institute analysis of carbapenem resistance reveals remarkable diversity, unexplained mechanisms, and limited clonal outbreaks. *Proc Natl Acad Sci U S A* 2017; **114**: 1135–40. https://doi.org/10.1073/pnas.1616248114

41 Kwong JC, Lane CR, Romanes F *et al*. Translating genomics into practice for real-time surveillance and response to carbapenemase-producing Enterobacteriaceae: evidence from a complex multi-institutional KPC outbreak. *PeerJ* 2018; **6**: e4210. https://doi.org/10.7717/peerj.4210

42 Evans DR, Griffith MP, Sundermann AJ *et al*. Systematic detection of horizontal gene transfer across genera among multidrug-resistant bacteria in a single hospital. *eLife* 2020; **9**: e53886. https://doi.org/10.7554/eLife.53886

43 Lerminiaux NA, Cameron ADS. Horizontal transfer of antibiotic resistance genes in clinical environments. *Can J Microbiol* 2019; **65**: 34–44. https://doi.org/10.1139/cjm-2018-0275

44 Stadler T, Meinel D, Aguilar-Bultet L *et al*. Transmission of ESBL-producing Enterobacteriaceae and their mobile genetic elements—identification of sources by whole genome sequencing: study protocol for an observational study in Switzerland. *BMJ Open* 2018; **8**: e021823. https://doi.org/10.1136/bmjopen-2018-021823