Genomic characterisation of multidrug-resistant *Escherichia coli*, *Klebsiella pneumoniae*, and *Acinetobacter baumannii* in two intensive care units in Hanoi, Viet Nam: a prospective observational cohort study

Leah W Roberts, Le Thi Hoi, Fahad A Khokhar, Nguyen Thi Hoa, Tran Van Giang, Cuong Bui, Tran Hai Ninh, Dao Xuan Co, Nguyen Gia Binh, Hoang Bao Long, Dang Thi Huong, James E Bryan, Archie Herrick, Theresa Feltwell, Behzad Nadjm, H Rogier van Doorn, Julian Parkhill, Nguyen Vu Trung, Nguyen Van Kinh, Zamin Iqbal, M Estée Török**

Summary

Background Viet Nam has high rates of antimicrobial resistance (AMR) but little capacity for genomic surveillance. This study used whole genome sequencing to examine the prevalence and transmission of three key AMR pathogens in two intensive care units (ICUs) in Hanoi, Viet Nam.

Methods A prospective surveillance study of all adults admitted to ICUs at the National Hospital for Tropical Diseases and Bach Mai Hospital was done between June 19, 2017, and Jan 16, 2018. Clinical and environmental samples were cultured on selective media, characterised with MALDI TOF mass spectrometry, and sequenced with Illumina. Phylogenies based on the de-novo assemblies (SPAdes) were constructed with MAFFT (PARsnp), Gubbins, and RAxML. Resistance genes were detected with Abricate against the US National Center for Biotechnology Information database.

Findings 3153 *Escherichia coli*, *Klebsiella pneumoniae*, and *Acinetobacter baumannii* isolates from 369 patients were analysed. Phylogenetic analysis revealed predominant lineages within *A baumannii* (global clone 2, sequence types ST2 and ST571) and *K pneumoniae* (ST15, ST16, ST656, ST11, and ST147) isolates. Isolation from stool was most common with *E coli* (87·0%) followed by *K pneumoniae* (62·5%). Of the *E coli*, 85·0% carried a $bla_{CTX\text{-}M}$ variant, while 81·8% of *K pneumoniae* isolates carried bla_{NDM} (54·4%), or bla_{KPC} (45·1%), or both. Transmission analysis with single nucleotide polymorphisms identified 167 clusters involving 251 (68%) of 369 patients, in some cases involving patients from both ICUs. There were no clear differences between the lineages or AMR genes recovered between the two ICUs.

Interpretation This study represents the largest prospective surveillance study of key AMR pathogens in Vietnamese ICUs. Clusters of closely related isolates in patients across both ICUs suggests recent transmission before ICU admission in other health-care settings or in the community.

Funding UK Medical Research Council Newton Fund, Viet Nam Ministry of Science and Technology, Wellcome Trust, Academy of Medical Sciences, Health Foundation, and UK National Institute for Health and Care Research Cambridge Biomedical Research Centre.

Introduction

Low-income and middle-income countries (LMICs) have reported widespread antimicrobial resistance (AMR) in health-care, community, and agricultural settings. In southeast Asia, dense human populations, intensive animal farming, unrestricted access to antibiotics, and scarce laboratory infrastructure have all contributed to the rapid expansion of AMR.[1,2] Much of this burden arises from excessive use of antimicrobials in human and animal populations.

In Viet Nam, antimicrobial usage has been estimated to be two times higher in people, and one and a half times higher in animals, compared with the EU.[3] Despite legal restrictions in Viet Nam, antibiotics are often dispensed without prescriptions in the community.[4] Broad-spectrum antibiotics are also commonly administered in health-care settings to mitigate the effects of the scarce capacity for microbiological testing and infection control.[4,5] Detection of both AMR bacteria and antimicrobials have been recorded in the environment,[6,7] hospital waste,[8] and food sources.[9]

Extensive AMR has led to increased pressure on hospitals and is particularly problematic in critical care settings. Although AMR surveillance based on phenotypic testing in Viet Nam has improved since 2015, the infrastructure required for systematic genomic surveillance remains to be established. Genomic analysis is important to identify circulating lineages; however,

Lancet Microbe 2022

Published Online October 4, 2022 https://doi.org/10.1016/S2666-5247(22)00181-1

*Contributed equally

European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Hinxton, UK (L W Roberts PhD, Z Iqbal DPhil); **Department of Medicine** (L W Roberts, F A Khokhar BSc, J E Bryan MBBS, A Herrick BSc MPH, T Feltwell MSc, M E Török PhD FRCP), **Cambridge Institute of Therapeutic Immunology and Infectious Diseases** (F A Khokhar), **and Department of Veterinary Medicine** (Prof J Parkhill PhD), **University of Cambridge, Cambridge, UK; National Hospital for Tropical Diseases, Hanoi, Viet Nam** (L T Hoi MD, N T Hoa MD, T V Giang MD, T H Ninh MD, D T Huong MD, N V Trung MD, N V Kinh MD); **Hanoi Medical University, Hanoi, Viet Nam** (L T Hoi, T V Giang, H B Long MD, D T Huong, N V Trung); **Department of Microbiology and National Tuberculosis Reference Laboratory, National Lung Hospital, Hanoi, Viet Nam** (N T Hoa); **Bach Mai Hospital, Hanoi, Viet Nam** (C Bui MD, D X Co MD, N G Binh MD); **Medical Research Council The Gambia at the London School of Hygiene & Tropical Medicine, Fajara, The Gambia** (B Nadjm MD FRCP); **Wellcome Trust Major Overseas Programme, Hanoi, Viet Nam** (B Nadjm, H R van Doorn MD); **Nuffield Department of Medicine, University of Oxford, Oxford, UK** (H R van Doorn); **Departments of Infectious Diseases and Microbiology,**

Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK (M E Török)

Correspondence to: M Estée Török, Department of Medicine, University of Cambridge, Cambridge CB2 0QQ, UK
et317@cam.ac.uk

Research in context

Evidence before this study

We searched PubMed using the search terms "Vietnam", "whole genome sequencing" and "hospital" between Jan 1, 2000, and Jan 1, 2017. This search retrieved only 20 publications, 14 of which were related to bacterial pathogens. From these 14, two focused on *Klebsiella pneumoniae*, three on *Acinetobacter baumannii*, and none on *Escherichia coli*. Neither of the *K pneumoniae* studies were based in Viet Nam, and they analysed only a modest number of strains (approximately 300 in one study and 90 in the other). Two of the *A baumannii* studies focused on Vietnamese intensive care settings but on only a small number of strains (147 in one study and 93 in the other) from earlier in the decade (2003–13). The remaining *A baumannii* study looked at in-vitro emergence of colistin resistance. None of the studies looked at broad contemporary circulation of antimicrobial resistance (AMR) in the three species in Viet Nam, or possible transmission chains on a large (>1000 strain) scale. This is a major gap, as *E coli, K pneumoniae*, and *A baumannii* were in the top five pathogens, each responsible for more than 250 000 deaths, associated with AMR in 2019, and in the top four pathogens associated with AMR-attributable deaths in southeast Asia.

Added value of this study

This is the largest prospective surveillance study of three key AMR pathogens (*E coli, K pneumoniae*, and *A baumannii*) done in critical care settings in Viet Nam. Sampling was restricted to patients who were colonised or infected with extended-spectrum β-lactamase-producing or carbapenem-resistant organisms. Colonisation with more than one organism was very common, with multidrug-resistant *E coli* being predominant in stool samples. A small number of predominant lineages were identified for *K pneumoniae* and *A baumannii*, whereas the *E coli* isolates were highly genetically diverse. Many genomic clusters were identified within the two intensive care units (ICUs), some of which spanned both ICUs. There were no significant differences between lineages or AMR genes between the two ICUs.

Implications of all the available evidence

This study found high rates of three key AMR pathogens in adults admitted to two ICUs in Viet Nam. Analysis of transmission broadly across the dataset identified similar lineages and AMR genes in both ICUs suggesting that dissemination occurred before ICU admission, either in referral hospitals or in community settings before hospital admission. Strategies to tackle AMR in Viet Nam will need to account for transmission and AMR dissemination before ICU admission by extending surveillance more widely across hospital and community settings.

LMICs remain relatively understudied, with few studies done in Vietnamese hospitals.[10–13]

To address this knowledge gap, we did a prospective genomic surveillance study of key AMR pathogens in two hospitals in Viet Nam. We focussed our analysis on the three most commonly isolated species (*Escherichia coli*, *Klebsiella pneumoniae*, and *Acinetobacter baumannii*) that were extended-spectrum β-lactamase (ESBL) producers or carbapenem resistant. We describe the genomic diversity of these isolates and broadly gauge the level of isolate relatedness to establish a conservative underestimate of transmission events during and immediately before intensive care unit (ICU) admission.

Methods

Study design, setting, and participants

See Online for appendix 1

This prospective observational cohort study was done in two hospitals, the National Hospital for Tropical Diseases (NHTD) and Bach Mai Hospital (BMH), in Hanoi, Viet Nam, between June 19, 2017, to Jan 16, 2018. All adult patients admitted to the ICUs of the two hospitals during the study period were eligible for inclusion in the study; there were no exclusion criteria. We targeted patients who were admitted to ICUs because we hypothesised that they would be the most likely to have been treated with antibiotics and to harbour AMR pathogens. Screening specimens (stool or rectal swabs, urine, skin or wound swabs or pus, and sputum or tracheal aspirates) were collected from ICU patients on admission, on discharge, and weekly during their ICU stay. Environmental samples were collected using flocked swabs (from door handles, bed rails, medical equipment, and patient tables) once per month.

The study protocol was approved by the Scientific and Ethical Committees of the National Hospital for Tropical Diseases and Bach Mai Hospital and by the University of Cambridge Human Biology and Research Ethics Committee (reference HBREC 2017.09). Written informed consent was obtained from the patient or from their relative before enrolment in the study.

Procedures

Specimens were cultured on selective media to identify ESBL producers and carbapenem-resistant organisms (appendix 1 pp 2–3). Target organisms (*E coli*, *A baumannii*, and *K pneumoniae*) were identified with MALDI-TOF mass spectrometry (Bruker Diagnostics, Bremen, Germany) and stored at –80°C before being shipped to the University of Cambridge, UK, where the samples underwent antimicrobial susceptibility testing (Vitek-2, BioMérieux, Marcy L'Étoile, France) and DNA extraction using QIACube and the QIAamp 96 DNA QIACube HT kit (Qiagen, Hilden, Germany).

DNA extracts were transferred to the Wellcome Sanger Institute, UK, for library preparation and sequencing;

DNA was sequenced in two batches on an Illumina HiSeq X10 machine (Illumina, San Diego, CA, USA). Further details of laboratory methods including read quality control, genome assembly, phylogenetic analysis, antibiotic resistance gene detection, multilocus sequence typing, and transmission cluster analysis are available in appendix 1 (pp 2–5).

Briefly, phylogenetic trees were built from core multi-alignments (PARsnp v1.2), filtered using Gubbins (v2.3.5) and constructed with RAxML (GTR-GAMMA model; v8.2.12).

Outcomes and analyses

We determined the number of *E coli*, *K pneumoniae*, and *A baumannii* isolates cultured from clinical and environmental samples collected during the study period. We did whole-genome sequencing of these isolates followed by phylogenetic analyses to examine genomic diversity and relatedness. We also determined the presence of antibiotic resistance genes. We did descriptive statistical analyses of numerical data.

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

A total of 3367 isolates were cultured, comprising *E coli* (n=765), *K pneumoniae* (n=1372), and *A baumannii* (n=1230). 31 isolates were excluded from the analysis because of poor assembly quality. A further 150 isolates were excluded because of suspected interspecies contamination, and 33 isolates were excluded because of suspected intraspecies (strain-level) contamination (appendix 1 p 14). Thus 3153 isolates (721 *E coli*, 1316 *K pneumoniae*, 1116 *A baumannii*), comprising 2891 isolates from 369 patients and 262 environmental isolates, passed quality filtering and were included in the final analyses.

Overall, a total of 1042 isolates (993 clinical and 49 environmental) were collected from BMH and 2111 isolates (1898 clinical and 213 environmental) from NHTD. The participant baseline characteristics and outcomes for each hospital are summarised in table 1 and appendix 1 (p 15). The number of samples collected was higher, and the median length of ICU stay was longer, at NHTD than BMH.

146 (40%) of 369 patients (55 at BMH and 91 at NHTD) were colonised or infected with all three bacterial species; 133 (36%) patients (66 at BMH and 67 at NHTD) with two of the three species; and 90 (24%) patients (61 at BMH and 29 at NHTD) had only one species detected.

Both *E coli* (627 [87·0%] of 721) and *K pneumoniae* (822 [62·5%] of 1316) were isolated primarily from stool or rectal swabs. *K pneumoniae* was also isolated from other sites including sputum (325 [24·7%] of 1316), urine samples (63 [4·8%] of 1316), and skin swabs (17 [1·3%] of 1316). By contrast, *A baumannii* isolates were predominantly isolated from sputum (621 [55·6%] of 1116), followed by stool and rectal swabs (247 [22·1%] of 1116), urine (49 [4·4%] of 1116), and skin swabs (36 [3·2%] of 1116). *A baumannii* also accounted for the highest number of environmental isolates (161 [14·4%] of 1116), compared with 85 (6·5%) of 1316 for *K pneumoniae* and 16 (2·2%) of 721 for *E coli*.

Phylogenetic trees for each species were constructed to explore lineage diversity. The *E coli* isolates were highly

	Participants from Bach Mai Hospital (n=182)	Participants from the National Hospital of Tropical Diseases (n=187)
Sex		
Male	104 (57%)	114 (61%)
Female	71 (39%)	50 (27%)
Not recorded	7 (4%)	23 (12%)
Median age, years	55 (42–66)	57·5 (45·75–69)
Age not recorded	5 (3%)	19 (10%)
Age by sex, years		
Male	55 (42–63)	56·5 (46·25–67·75)
Female	57 (35–68)	59·5 (40·25–69)
Male age groups		
0–20 years	3	5
21–40 years	19	19
41–60 years	48	42
61–80 years	28	41
81–100 years	6	7
Female age groups		
0–20 years	3	3
21–40 years	19	10
41–60 years	21	13
61–80 years	23	20
81–100 years	5	4
Duration of ICU admission		
Length of stay, days	8 (6)	21 (16)
Length of stay not recorded	4 (2%)	7 (4%)
Outcome at ICU discharge		
Stable, discharged home	10 (6%)	38 (20%)
Improved, transferred to another ward	117 (64%)	83 (44%)
Deteriorated, transferred to another ward	3 (2%)	7 (4%)
Discharged home to die	44 (24%)	43 (23%)
Died in hospital	4 (2%)	9 (5%)
Not recorded	4 (2%)	7 (4%)

Data are n (%), median (IQR), n, or mean (SD). ICU=intensive care unit.

Table 1: **Summary of study participants**

diverse, with isolates spread over eight phylogroups and 80 sequence types (figure 1). The most prevalent sequence type was ST648 (phylogroup A; 11·8%), followed by ST410 (phylogroup C; 9·7%), ST617 (phylogroup A; 9·2%), ST131 (phylogroup B2; 7·9%), and ST1193 (phylogroup B2; 7·4%). Overall, 33 of 80 sequence types had only one representative isolate in this dataset.

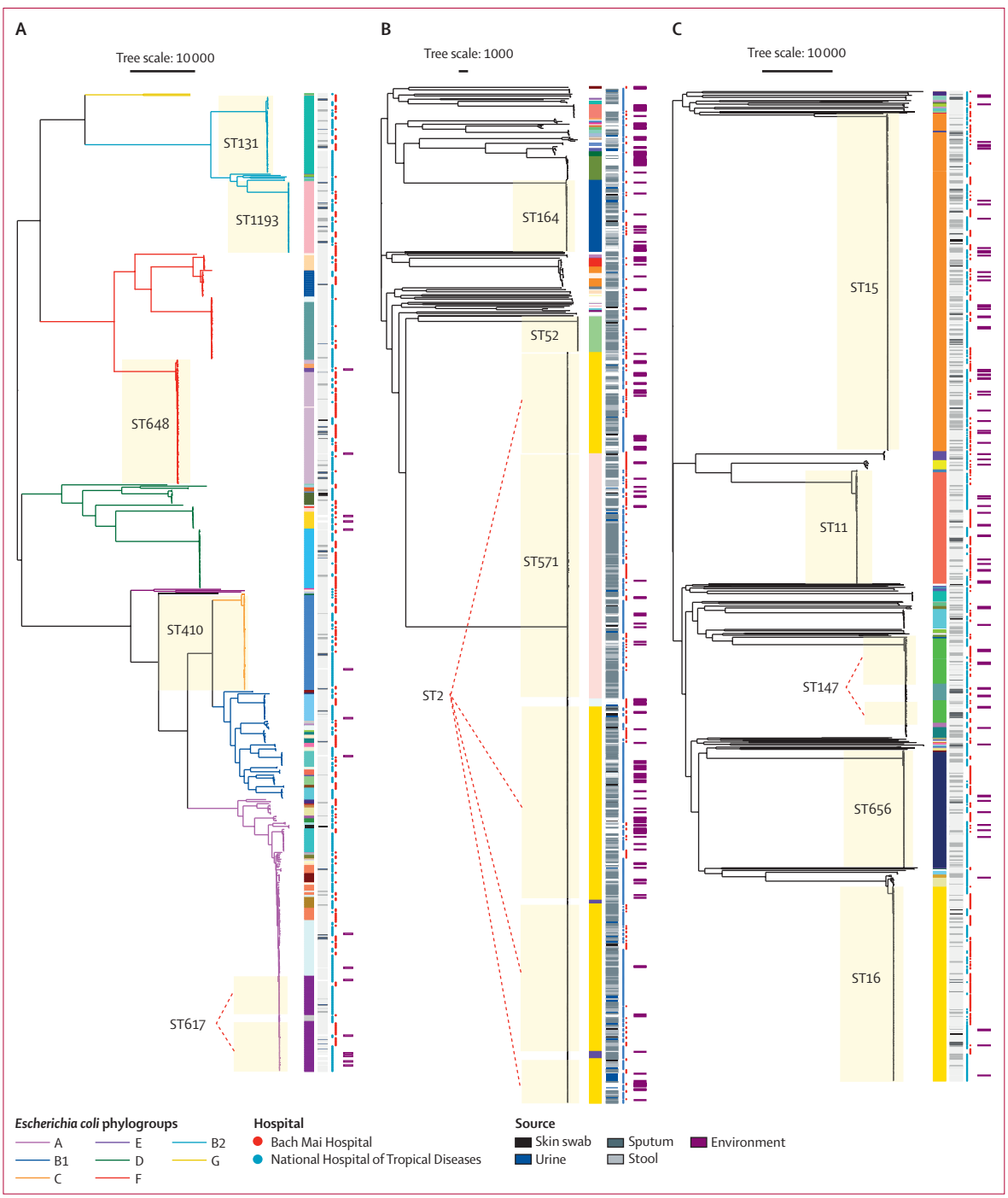

Figure 1: **Whole genome phylogenies**
(A) *Escherichia coli*. (B) *Acinetobacter baumannii*. (C) *Klebsiella pneumoniae*. Recombination-filtered core-single nucleotide polymorphism trees with midpoint root. Tree metadata includes (from left to right column beside trees): multilocus sequence typing, source, and hospital. Outermost purple bars indicate environmental isolates. Branches corresponding to *E coli* phylogroups are coloured accordingly. Main sequence types are highlighted in the image with yellow boxes. Tree scale is number of substitutions.

By contrast, the *K pneumoniae* and *A baumannii* isolates had a small number of dominant lineages. 1094 (83%) of the 1316 *K pneumoniae* isolates were from one of five sequence types, including ST15 (445 [34%]), ST16 (259 [20%]), ST656 (154 [12%]), ST11 (148 [11%]), and ST147 (88 [7%]). The majority of the 1116 *A baumannii* isolates were global clone 2 (832 [75%])[14] and mainly belonged to ST2 (536 [48%]) and ST571 (269 [24%]; based on the Pasteur scheme). We did not identify any specific relationship between sequence types and hospitals, with all of the major sequence type lineages represented in both ICUs.

To gain broader insight into the lineages, we selected globally representative strains to contextualise our dataset. Addition of these global representatives into the *E coli* phylogeny showed that most isolates belonged to a globally diverse set of sequence types that were not unique to Viet Nam, but found across parts of North America, Europe, and Asia (appendix 1 p 16). Similarly, several of the major *K pneumoniae* lineages were represented globally, particularly ST147, ST11 (mainly from China and the USA), and ST15 (mainly other Asian countries; appendix 1 p 17). However, it was also clear that local expansion was prominent, particularly among the ST656, ST16, and ST15 lineages. For *A baumannii*, we focused primarily on global clone 2 isolates (appendix 1 p 18). There was very little representation of global strains within our dataset, and those that were available consisted mainly of strains from other parts of Asia. Closer inspection of all global representatives found several strains in each species that were closely related (less than five core single nucleotide polymorphisms [SNPs]) to isolates in our dataset (appendix 1 pp 6, 8).

Almost all isolates carried acquired resistance genes to at least three antibiotic classes, with 649 (90%) of *E coli*, 1134 (86%) of *K pneumoniae*, and 452 (41%) of *A baumannii* isolates carrying genes across five antibiotic classes (appendix 1 p 19). There were no discernible differences based on sample source or hospital, with the exception of *E coli* detected in wound and skin swabs (n=6), which appeared on average to carry resistance to more antibiotic classes. *K pneumoniae* isolates tended to fall into one of three high density regions (appendix 1 p 19) due to lineage-specific carriage of acquired resistance genes (ST15 isolates tended to carry resistance to more classes, whereas ST16 often carried the least).

Resistance to antibiotics classes varied across the *E coli* phylogeny, reflective of the diversity of strains within the dataset (appendix 1 p 20). Bla_{CTX-M} genes were found in 613 of the 721 *E coli* isolates (table 2), with $bla_{CTX-M-15}$ (259 [36%]), $bla_{CTX-M-27}$ (213 [30%]), and $bla_{CTX-M-55}$ (119 [17%]) being the most prevalent (appendix 1 p 9). Bla_{KPC-2} (94 [13%]) and $bla_{NDM-[1,4,5,7]}$ (173 [24%]) were present sporadically across the phylogroups, suggesting independent acquisition events. Only 28 (4%) isolates carried $mcr_{[1.1,3.5,8.2,9.1]}$ genes conferring resistance to colistin. Again, these seemed to be independent acquisitions, except for an ST206 cluster (phylogroup A; n=11) involving three patients from NHTD.

Conversely, multidrug-resistant gene presence across the *K pneumoniae* isolates were consistent with the main lineages, suggesting clonal expansion rather than diverse sampling of the species (appendix 1 p 21). Similar to the *E coli*, incidence of $bla_{CTX-M-15}$ (493 [37·5%] of 1316) was high, but less so than bla_{KPC-2} (593 [45·1%]) and bla_{NDM} (716 [54·4%]; table 2).

Acquired AMR genes were overall less prevalent among the *A baumannii* isolates. Similar to the *K pneumoniae*, resistance to specific classes tended to be a feature of each distinct lineage, suggesting clonal expansion (appendix 1 p 22). The carbapenemase gene bla_{OXA-23} was present in 927 (83%) of 1116 isolates, with bla_{OXA-58} (64 [5%]) and bla_{OXA-72} (3 [<1%]) present at much lower frequencies

	Escherichia coli isolates (n=721)	*Acinetobacter baumannii* isolates (n=1116)	*Klebsiella pneumoniae* isolates (n=1316)
Tetracycline	563 (78·1%)	701 (62·8%)	753 (57·2%)
Sulphonamide	649 (90·0%)	746 (66·8%)	969 (73·6%)
Fluoroquinolone	161 (22·3%)	19 (1·7%)	1187* (90·2%)
Colistin	28 (3·9%)	0	10 (0·8%)
Fosfomycin	48 (6·7%)	5 (0·4%)	1316* (100%)
Macrolide, lincosamide, and streptogramin	579 (80·3%)	816 (73·1%)	623 (47·3%)
Trimethoprim	621 (86·1%)	41 (3·7%)	982 (74·6%)
Phenicols	274 (38·0%)	176 (15·8%)	675 (51·3%)
Bleomycin	179 (24·8%)	36 (3·2%)	717 (54·5%)
β-lactamase	718 (99·6%)	1018 (91·2%)	1286 (97·7%)
Class C			
EC	721† (100%)	2 (0·2%)	0
ACT	0	1 (0·1%)	0
CMY	209 (29·0%)	1 (0·1%)	2 (0·2%)
DHA	26 (3·6%)	2 (0·2%)	24 (1·8%)
Class A			
LAP	17 (2·4%)	0	142 (10·8%)
CARB	0	63 (5·6%)	0
PER	0	68 (6·1%)	0
TEM	347 (48·1%)	697 (62·5%)	686 (52·1%)
SHV	7 (1·0%)	9 (0·8%)	1292* (98·2%)
VEB	0	12 (1·1%)	3 (0·2%)
CTX	613 (85·0%)	4 (0·4%)	681 (51·8%)
KPC	94 (13·0%)	3 (0·3%)	593 (45·1%)
Class D			
OXA	251 (34·8%)	996‡ (89·2%)	611 (46·4%)
Class B			
IMP	0	6 (0·5%)	1 (0·1%)
NDM	173 (24·0%)	35 (3·1%)	716 (54·4%)
Rifamycin	114 (15·8%)	63 (5·6%)	810 (61·6%)
Aminoglycoside	680 (94·3%)	1115 (99·9%)	1290 (98·0%)
Streptothricin	11 (1·5%)	8 (0·7%)	0

*_fosA_ (fosfomycin), _oqxAB_ (fluoroquinolone), and bla_{SHV} intrinsic in *K pneumoniae*. †bla_{EC} intrinsic in *E. coli*. ‡bla_{ADC} and bla_{OXA} intrinsic in *A baumannii* (except OXA-[1,10,23,58,72]).

Table 2: **Summary of resistance genes found in the three bacterial species**

(appendix 1 p 9). The aminoglycoside resistance gene *armA* was also highly prevalent (844 [76%] of 1116 isolates).

Overall, 133 different AMR genes were detected in BMH and 154 were detected in NHTD. 49 genes were unique to either hospital (35 in NHTD, 14 in BMH), but were only detected at a prevalence of less than 0·1%, suggesting sporadic cases. The remaining 119 genes were the same across both hospitals. Genes with at least 1% prevalence in each hospital were almost identical. There were only five exceptions in which the gene prevalence was more than 1% in BMH but less than 1% in NHTD (bla_{NDM-4} [0·98%], *dfrA12* [0·92%], *rmtB1* [0·86%], *qnrB6* [0·74%], and $bla_{OXA-181}$ [0·56%]). We attempted to discover possible mobile genetic-element-driven AMR transmission by monitoring the fluctuation of the prevalence of AMR genes over time; however, we were unable to detect any substantial changes (appendix 1 pp 6, 23–24).

Most patients with *E coli* and *K pneumoniae* had isolates only from stool (218 [81%] of 270 patients with *E coli*, and 158 [54%] of 294 patients with *K pneumoniae*). Conversely, most patients with *A baumannii* (n=230) were detected only in sputum (86 [37%]), or in sputum and stool (65 [28%]). Repeat isolation of the same bacterial species on at least two separate occasions was common (*E coli* 138 [51%] of 270, *K pneumoniae* 169 [57%] of 294, and *A baumannii* 132 [58%] of 230). Of these patients, 60–70% had different sequence types (92 [67%] of 138 *E coli*, 115 [68%] of 169 *K pneumoniae*, 83 [63%] of 132 *A baumannii*; figure 2). A large proportion also had the same sequence type isolated at two or more timepoints (103 [75%] of 138 *E coli*, 149 [88%] of 169 *K pneumoniae*, and 123 [93%] of 132 *A baumannii*). For the 103 patients with the same *E coli* sequence type isolated over multiple timepoints, the majority (73 [70%]) were detected only from stool, with fewer cases found in both stool and sputum (14 [14%]) or stool and urine (10 [10%]). 45% of patients (n=67) with the same *K pneumoniae* sequence type were largely isolated from stool (53 [36%]). Similarly, 46% of patients (n=56) with the same *A baumannii* sequence type were isolated only from sputum (42 [34%]; appendix 1 p 12).

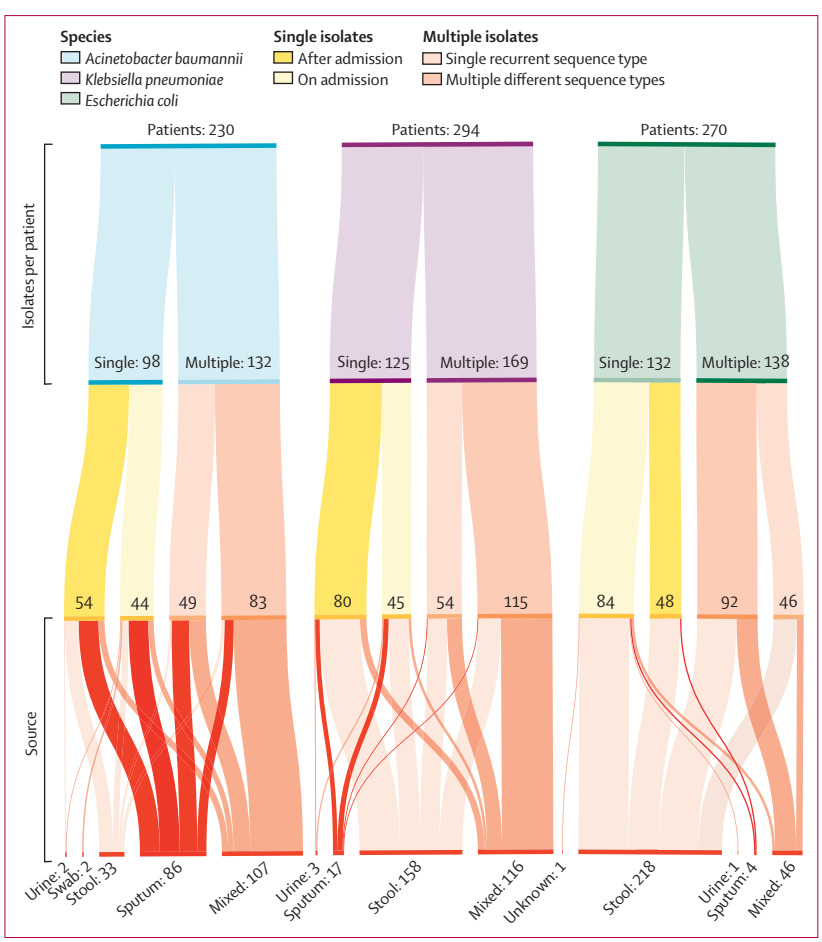

Figure 2: **Overview of strain diversity, recurrence, and source among study patients**
Patients refers to the total number of patients in this study that had at least one isolate of that species. Within each species, we evaluated whether patients had only a single isolate for that species, or multiple isolates. If only a single isolate, we identified whether it was collected on admission to the intensive care unit or after. For multiple isolates, we identified whether the patient's isolates were the same sequence type (recurrent) or a different sequence type. We finally looked at how many patients had isolates from one site (urine, swab, stool, or sputum) or mixed sites (any combination of sites).

Temporal analysis of the isolates found no obvious association of any timepoint with any sequence type to suggest an outbreak of a specific lineage. To identify closely related strains that could indicate recent transmission, we evaluated clusters on the basis of SNP distances across the core genome of each species for this dataset. Given the short sampling period, none of the three major species was likely to acquire more than one SNP while in the hospital. As such, we looked at samples with genomic evidence of most recent transmission: zero SNP clusters (followed up with higher thresholds). We chose this SNP threshold to broadly gauge the level of isolate relatedness and determine a conservative underestimate of transmission. Clusters were defined when they involved more than one patient. Clusters involving a single patient and environmental samples were not included.

Despite our conservative threshold, we identified several clusters. Most clusters were detected in *K pneumoniae* isolates (71 clusters, representing 38% of total isolates) and *A baumannii* isolates (74 clusters, representing 52% of total isolates; figure 3). *K pneumoniae* had some of the largest clusters, ranging in size from two to 79 isolates, whereas *A baumannii* clusters were smaller, from two to 33 isolates. Only 22 clusters were detected in *E coli* and were generally small (median three isolates [IQR 2–5·75]), representing only 13% of the *E coli* dataset. The sequence types with the greatest number of clusters in each species were ST410 (*E coli*), ST2 (*A baumannii*), and ST15 (*K pneumoniae*; appendix 1 p 11).

For all three species, the majority of clusters (119 [71%] of 167) were detected in patients within a single ICU. Evaluating admission and discharge dates further confirmed patient overlap in these clusters (figure 3, appendix 1 pp 25–30). Patients involved in zero SNP

clusters overlapped with more patients during their time in the ICU (mean 29 patients [SD 15]) than patients not involved in clusters (20 patients [SD 12]). Compared with *E coli* and *K pneumoniae*, *A baumannii* clusters were more often associated with environmental isolates (24 [32%] of 74 clusters; figure 3). 15 (21%) of 71 *K pneumoniae* clusters and only two (9%) of 22 *E coli* clusters were associated with environmental isolates. *K pneumoniae* environmental isolates were more often found in within-hospital (ie, ICU) clusters (n=11) compared with between-hospital clusters (n=4).

In addition to suspected within-ICU transmission, we also detected several clusters involving patients from both hospital ICUs (figure 3). The most pronounced example of this was a large ST15 *K pneumoniae* cluster involving 79 isolates from 38 patients and six environmental samples (appendix 1 p 31).

The identification of closely related isolates between independently operating ICUs suggested that there might have been a common source located outside the ICU—eg, admission to the same shared location (such as a ward in the same hospital as the current ICU, or the same ward in a different prior hospital) before admission to ICU. To find out whether particular lineages were associated with acquisition within the ICU, we assessed diversity on arrival (ie, admission samples) versus diversity within the ICU (all other samples). 275 (75%) of 369 patients had a positive culture on admission. Based on sequence type alone, we found a slight increase in diversity in ICUs versus on arrival (appendix 1 p 32). However, the unique sequence types recovered in either setting represented only a small portion of the isolates overall. All the main lineages for each species were found on both admission and within ICUs (appendix 1 pp 32–33). In relation to our zero SNP clusters, we found that clusters involving both ICUs more often had a patient with an admission positive sample (*A baumannii*: 24 [89%] of 27 clusters; *K pneumoniae:* 13 [76%] of 17) than clusters within a single ICU (*A baumannii*: 24 [53%] of 45; *K pneumoniae:* 25 [47%] of 53; appendix 1 p 12). For *E coli*, there was no bias of admission-positive isolates in either cluster group. Overall, at least half of the clusters

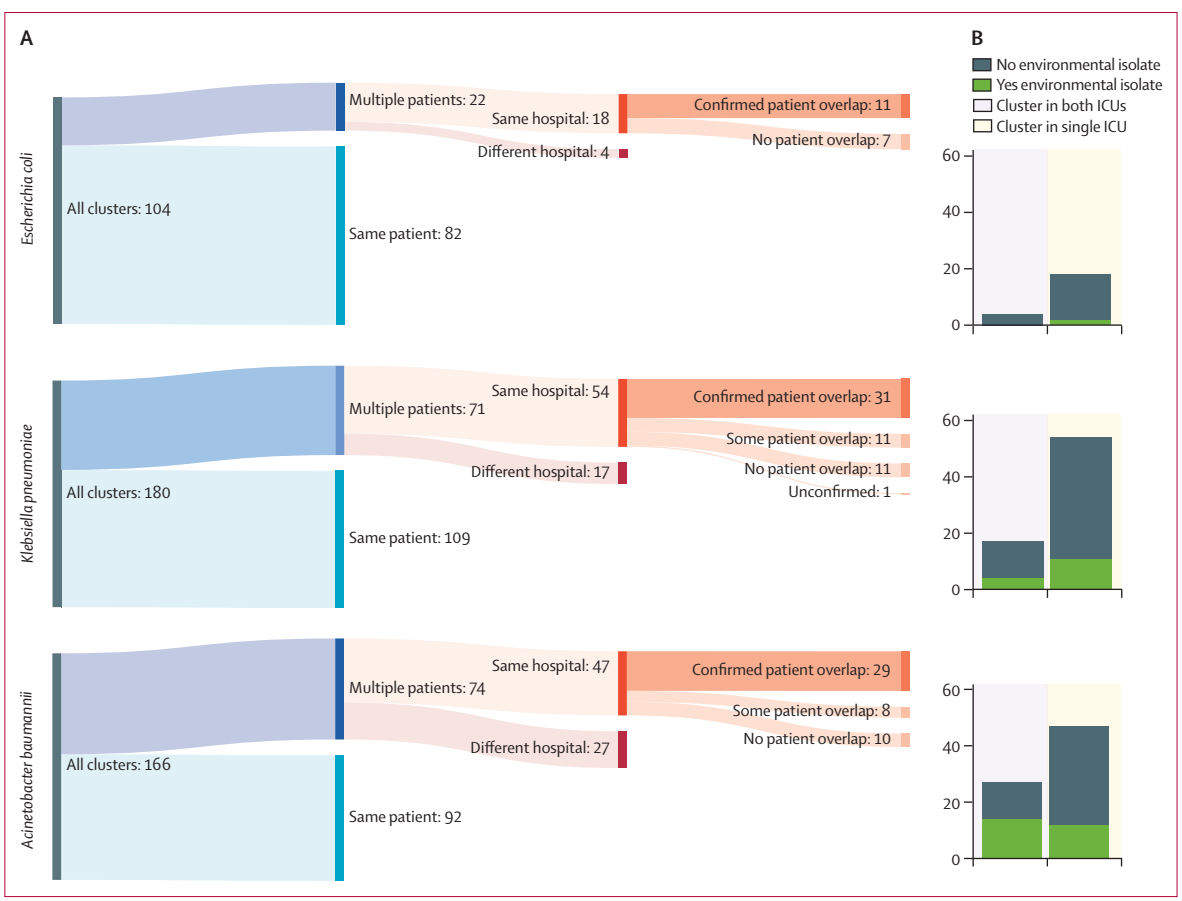

Figure 3: **Summary of zero SNP clusters in all species**
(A) Clusters were defined as multiple patients (samples were derived from at least two different patients) or same patient (isolates were derived from the same patient, or only a single patient and the environment). Epidemiological evidence to support clusters was defined as confirmed patient overlap (all patient ICU stays overlap with another in the same cluster), some patient overlap (at least two patient ICU stays overlap), and zero patient overlap between all patients in cluster. (B) Environmental isolates in clusters were counted if an environmental isolate was found in that cluster. SNP=single nucleotide polymorphism. ICU=intensive care unit.

for all species involved an isolate collected from an admission sample (*E coli*: 11 [50%] of 22; *A baumannii*: 48 [67%] of 72; *K pneumoniae: 38* [54%] of 70).

Over the course of this study, 251 patients (representing 68% of the cohort) were involved in 167 clusters across the three species. 112 (45%) patients were involved in only a single cluster during their time in the ICUs (figure 4). However, the remaining 139 (55%) patients were involved in at least two clusters, with one (<1%) patient involved in 12 clusters. For the 139 patients with at least two clusters, 20 (14%) had clusters from all three species, 94 (68%) had clusters from two species, and 25 (18%) had only one species. Overall, we saw a general trend towards more clusters in a single patient as they spent more time in the ICU ward.

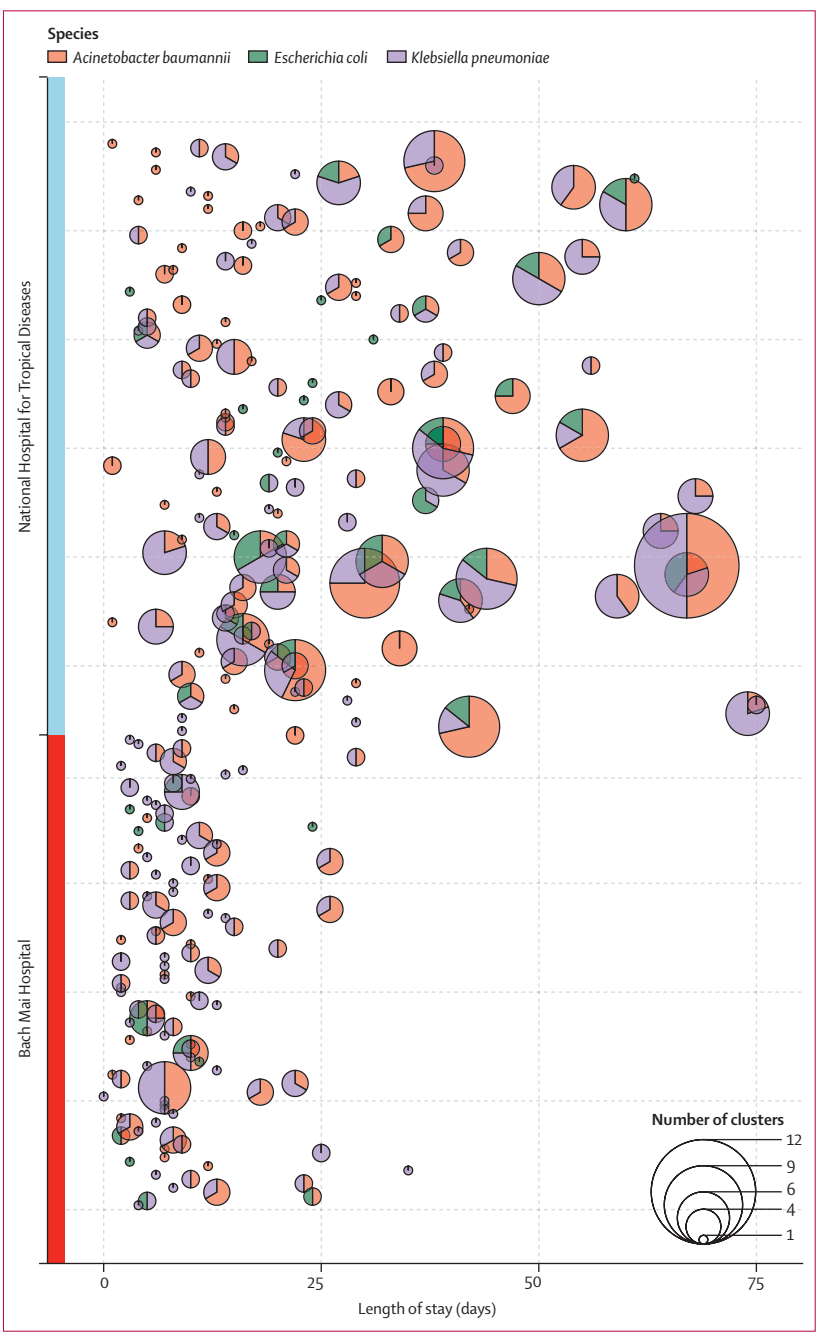

Figure 4: **Scatterpie showing the number of clusters in patients across all species**
Y-axis represents patients from the Bach Mai Hospital or the National Hospital for Tropical Diseases involved in at least one zero SNP cluster. One pie is plotted per patient at the duration of their stay. Each circle represents zero SNP clusters in a single patient. The size of the circle corresponds to the number of clusters, and the colour corresponds to the species. SNP=single nucleotide polymorphism.

To find out if any of our (zero SNP) clusters were potentially derived from a single original cluster predating their time in the ICU, we looked at SNP distances between clusters of the same sequence type (appendix 1 p 34). At a threshold of five SNPs, several of the prominent sequence types within each species formed large clusters, including ST804 in *A baumannii* and ST16 in *K pneumoniae*. At this threshold, we found 29 clusters in the *A baumannii* dataset (originally 74), 23 clusters in the *K pneumoniae* (originally 71) and 19 clusters in *E coli* (originally 22). Using these larger five SNP clusters, we found that slightly more patients (123 *vs* 112 previously) might have been involved in only a single cluster during their stay (appendix 1 p 35). 128 patients had two clusters, with the maximum number of clusters in a single patient being seven (n=3 patients).

Discussion

Here we present a large prospective genomic surveillance study of ESBL-producing or carbapenemase-producing organisms from three key AMR pathogens from two hospital ICUs in Viet Nam. Despite limiting our study samples to ESBL-producing or carbapenem-resistant isolates belonging to three species, we identified many isolates (an average of 17 isolates per day from patients in BMH and ten per day in NHTD). Similar settings in England, UK, have found a small incidence of ESBL-producing or carbapenemase-producing Enterobacterales.[15,16] However, across Europe and the USA, the spread of endemic ESBL-producing and carbapenemase-producing Enterobacterales has increased,[17] particularly *E coli* and *K pneumoniae*,[18] with extensive spread of high-risk lineages, such as *K pneumoniae* ST258 and ST512 in the USA,[19] Israel, Italy,[20] and Greece.[21] Meanwhile, the burden of carbapenem-resistant *A baumannii* remains higher in south Asia than in high-income regions.[22]

Several of the dominant circulating lineages described here have been identified previously. All the predominant *E coli* sequence types (ST617, ST648, ST410, ST131, and ST1193) are well known global lineages.[23] *K pneumoniae* ST15 is widespread in Asia and parts of Europe, but mostly absent in the Americas where ST258 and ST512 are more common.[24] ST11 is prevalent in China but has been reported previously in Viet Nam.[25] *A baumannii* global clone 2 has caused multiple outbreaks of carbapenem-resistant *A baumannii* globally,[26] including in Viet Nam,[27] and has displaced *Pseudomonas aeruginosa* as the main causative agent of ventilator-associated pneumonia in this region.[28]

In this study, we aimed to measure the breadth and volume of circulating ESBL-producing and carbapenem-resistant isolates from three species of bacteria in two ICUs. Rather than finding and investigating specific outbreaks, we set out to broadly estimate the number of transmissions occurring across the dataset (encompassing both within the ICU and the few months before admission), and show how they were distributed across the captured species diversity. Even when we used exact pairwise identity within the core genome, we found that 68% of patients were involved in recent transmission of several sequence types with no dominant spreading lineage. Just looking at these very closely related pairs alone provided a picture of high levels of circulating AMR in a wide range of lineages.

Our observation of clusters involving patients from both hospitals could be explained by a source outside ICUs, including other wards or hospitals that might have referred patients to intensive care,[29] or AMR strains that have been acquired in the community,[30] reflecting high rates of antibiotic use. Based on the similarity between lineages and AMR across both ICUs, we suggest that transmission is mainly circulating outside ICUs, where it is then further propagated. In this context, we should potentially reconsider AMR surveillance and control. Knowing whether hospitals are the primary source of AMR bacteria (and subsequent transmission into the community),[18] or whether the high rates are part of a more general, endemic pattern of circulating resistant strains is important for our understanding of local transmission dynamics. If the latter is true, this leads to very different national and indeed global management plans.

We acknowledge that whole genome sequencing of bacteria gives little temporal resolution on transmission, particularly with patients in an ICU for just 3 months. Whole genome sequencing to identify transmission events is more feasible in species with high mutation rates (eg, some viruses) and in low-AMR settings where community transmission is scarce and epidemiological surveillance can focus on specific lineages with detailed metadata. In our setting, the level of AMR circulation was too high to provide detailed transmission analyses (even with epidemiological support) over a large and diverse sample set. We were also limited by the single colony pick method; some transmission isolates were likely to be missed, and we could not estimate the true level of within-patient diversity. Similarly, we expect a patient's flora to contain some level of natural diversity, which is missed with our conservative SNP threshold and will inflate the number of within-patient clusters.

We also acknowledge other limitations to our study. First, we did not explore plasmid profiles among the samples because of the limitations of short read sequence data. This restricted our ability to detect interspecies and intraspecies transmission of AMR genes via plasmids. Second, we looked only at acquired AMR genes, and did not investigate resistance caused by point mutations. As such, we may have underestimated the resistance profiles of some isolates that were resistant via point mutations. Finally, this study focused on patient and environmental samples only. Therefore, we were unable to investigate potential transmission events involving hospital staff or visitors.

Nevertheless, we present the largest prospective surveillance study to date of multidrug-resistant *E coli*, *A baumannii*, and *K pneumoniae* in patients in critical care in Viet Nam, revealing frequent transmission of highly resistant bacteria within and between two ICU settings. Further work is required to expand genomic surveillance in hospital and community settings to inform AMR control strategies in Viet Nam.

Contributors

LWR, JP, NVK, ZI, and MET contributed to the conceptualisation of the manuscript. FAK, NTH, LTH, NGB, DXC, NTH, TVG, CB, THN, BN, HRvD, and NVT contributed to the data collection. NTH, FAK, JEB, AH, and TF contributed to the sample processing. LWR and ZI contributed to the methodology and formal analysis. LWR contributed to the writing (original draft). ZI and MET contributed to the writing (review and editing). NVT, HRvD, NVK, ZI, and MET contributed to the supervision of the study. LTH, NVT, HRvD, and MET contributed to the project administration. MET, NVK, and JP contributed to the funding acquisition. LWR and FAK accessed and verified all the data in this study. All authors had access to the data presented in this study and had final responsibility for the decision to submit for publication.

Declaration of interests

LWR reports travel fees from European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), and is on the Microbial Genomics Early Career Microbiologists board of reviewers, outside of the submitted work. HRvD is a board member of the Surveillance and Epidemiology of Drug Resistant Infections Consortium, outside of the submitted work. JP reports grants from the National Institute for Health and Care Research and Wellcome Trust and consulting fees and stock options from Next Gen Diagnostics, outside of the submitted work. ZI reports grants from Global Challenges Research Fund, UK Research and Innovation, and National Institute for Health and Care Research Health Protection Research Unit, travel fees from the US Centers for Disease Control and Prevention, and is part of the advisory board for Laboratory of Molecular Infection Medicine Sweden and CLIMB-BIG-DATA, outside of the submitted work. MET reports book royalties from Oxford University Press and teaching honoraria from Wellcome Sanger Institute, outside of the submitted work. All other authors declare no competing interests.

Data sharing

Genome sequence data have been deposited in the European Nucleotide Archive under the Bioproject PRJEB29424. A list of the sample accession numbers is available in appendix 2. Isolate genome assemblies (heterogenous sites masked and unmasked) are available on Figshare under the following DOI: 10.6084/m9.figshare.13303253 and 10.6084/m9.figshare.13302728.

See Online for appendix 2

Acknowledgments

We thank the patients for participating in this study and the clinical and laboratory staff of the National Hospital for Tropical Diseases and Bach Mai Hospital for their assistance with this study. We also acknowledge the sequencing team at the Wellcome Sanger Institute for their assistance with sequencing the samples including in the study. This study was funded by the UK Medical Research Council Newton Fund (grant MR/N029399/1), the Ministry of Science and Technology, Viet Nam (grant HNQT/SPĐP/04.16), and the Wellcome Trust (grant 206194). LWR is supported by a biomedical postdoctoral research fellowship by the European Molecular Biology Laboratory's European Bioinformatics Institute. MET was supported by a Clinician Scientist

Fellowship (funded by the Academy of Medical Sciences and the Health Foundation) and the National Institutes of Health Research Cambridge Biomedical Research Centre.

References

1 Zellweger RM, Carrique-Mas J, Limmathurotsakul D, et al. A current perspective on antimicrobial resistance in southeast Asia. *J Antimicrob Chemother* 2017; **72:** 2963–72.

2 Walther BA, Boëte C, Binot A, et al. Biodiversity and health: lessons and recommendations from an interdisciplinary conference to advise southeast Asian research, society and policy. *Infect Genet Evol* 2016; **40:** 29–46.

3 Carrique-Mas JJ, Choisy M, Van Cuong N, Thwaites G, Baker S. An estimation of total antimicrobial usage in humans and animals in Vietnam. *Antimicrob Resist Infect Control* 2020; **9:** 16.

4 Nguyen KV, Thi Do NT, Chandna A, et al. Antibiotic use and resistance in emerging economies: a situation analysis for Viet Nam. *BMC Public Health* 2013; **13:** 1158.

5 de Kraker ME, Stewardson AJ, Harbarth S. Will 10 million people die a year due to antimicrobial resistance by 2050? *PLoS Med* 2016; **13:** e1002184.

6 Tran NH, Hoang L, Nghiem LD, et al. Occurrence and risk assessment of multiple classes of antibiotics in urban canals and lakes in Hanoi, Vietnam. *Sci Total Environ* 2019; **692:** 157–74.

7 Yamasaki S, Le TD, Vien MQ, Van Dang C, Yamamoto Y. Prevalence of extended-spectrum β-lactamase-producing *Escherichia coli* and residual antimicrobials in the environment in Vietnam. *Anim Health Res Rev* 2017; **18:** 128–35.

8 Duong HA, Pham NH, Nguyen HT, et al. Occurrence, fate and antibiotic resistance of fluoroquinolone antibacterials in hospital wastewaters in Hanoi, Vietnam. *Chemosphere* 2008; **72:** 968–73.

9 Thapa SP, Shrestha S, Anal AK. Addressing the antibiotic resistance and improving the food safety in food supply chain (farm-to-fork) in southeast Asia. *Food Control* 2020; **108:** 106809.

10 Santona A, Taviani E, Hoang HM, et al. Emergence of unusual vanA/vanB$_2$ genotype in a highly mutated vanB$_2$-vancomycin-resistant hospital-associated *E faecium* background in Vietnam. *Int J Antimicrob Agents* 2018; **52:** 586–92.

11 Chung The H, Karkey A, Pham Thanh D, et al. A high-resolution genomic analysis of multidrug-resistant hospital outbreaks of *Klebsiella pneumoniae*. *EMBO Mol Med* 2015; **7:** 227–39.

12 Berglund B, Hoang NTB, Tärnberg M, et al. Colistin- and carbapenem-resistant *Klebsiella pneumoniae* carrying mcr-1 and blaOXA-48 isolated at a paediatric hospital in Vietnam. *J Antimicrob Chemother* 2018; **73:** 1100–02.

13 Berglund B, Hoang NTB, Tärnberg M, et al. Molecular and phenotypic characterization of clinical isolates belonging to a KPC-2-producing strain of ST15 *Klebsiella pneumoniae* from a Vietnamese pediatric hospital. *Antimicrob Resist Infect Control* 2019; **8:** 156.

14 Turton JF, Woodford N, Glover J, Yarde S, Kaufmann ME, Pitt TL. Identification of *Acinetobacter baumannii* by detection of the blaOXA-51-like carbapenemase gene intrinsic to this species. *J Clin Microbiol* 2006; **44:** 2974–76.

15 Trepanier P, Mallard K, Meunier D, et al. Carbapenemase-producing Enterobacteriaceae in the UK: a national study (EuSCAPE-UK) on prevalence, incidence, laboratory detection methods and infection control measures. *J Antimicrob Chemother* 2017; **72:** 596–603.

16 Wilson HJ, Khokhar F, Enoch DA, et al. Point-prevalence survey of carbapenemase-producing Enterobacteriaceae and vancomycin-resistant enterococci in adult inpatients in a university teaching hospital in the UK. *J Hosp Infect* 2018; **100:** 35–39.

17 Kazmierczak KM, de Jonge BLM, Stone GG, Sahm DF. Longitudinal analysis of ESBL and carbapenemase carriage among Enterobacterales and *Pseudomonas aeruginosa* isolates collected in Europe as part of the International Network for Optimal Resistance Monitoring (INFORM) global surveillance programme, 2013–17. *J Antimicrob Chemother* 2020; **75:** 1165–73.

18 David S, Reuter S, Harris SR, et al. Epidemic of carbapenem-resistant *Klebsiella pneumoniae* in Europe is driven by nosocomial spread. *Nat Microbiol* 2019; **4:** 1919–29.

19 Marsh JW, Mustapha MM, Griffith MP, et al. Evolution of outbreak-causing carbapenem-resistant *Klebsiella pneumoniae* ST258 at a tertiary care hospital over 8 years. *mBio* 2019; **10:** e01945–19.

20 Munoz-Price LS, Poirel L, Bonomo RA, et al. Clinical epidemiology of the global expansion of *Klebsiella pneumoniae* carbapenemases. *Lancet Infect Dis* 2013; **13:** 785–96.

21 Giakoupi P, Maltezou H, Polemis M, Pappa O, Saroglou G, Vatopoulos A. KPC-2-producing *Klebsiella pneumoniae* infections in Greek hospitals are mainly due to a hyperepidemic clone. *Euro Surveill* 2009; **14:** 19218.

22 Antimicrobial Resistance Collaborators. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet* 2022; **399:** 629–55.

23 Manges AR, Geum HM, Guo A, Edens TJ, Fibke CD, Pitout JDD. Global Extraintestinal Pathogenic *Escherichia coli* (ExPEC) lineages. *Clin Microbiol Rev* 2019; **32:** e00135–18.

24 Wyres KL, Lam MMC, Holt KE. Population genomics of *Klebsiella pneumoniae*. *Nat Rev Microbiol* 2020; **18:** 344–59.

25 Linh TD, Thu NH, Shibayama K, et al. Expansion of KPC-producing Enterobacterales in four large hospitals in Hanoi, Vietnam. *J Glob Antimicrob Resist* 2021; **27:** 200–11.

26 Hamidian M, Nigro SJ. Emergence, molecular mechanisms and global spread of carbapenem-resistant *Acinetobacter baumannii*. *Microb Genom* 2019; **5:** e000306.

27 Tada T, Miyoshi-Akiyama T, Shimada K, et al. Dissemination of clonal complex 2 *Acinetobacter baumannii* strains co-producing carbapenemases and 16S rRNA methylase ArmA in Vietnam. *BMC Infect Dis* 2015; **15:** 433.

28 Nhu NTK, Lan NPH, Campbell JI, et al. Emergence of carbapenem-resistant *Acinetobacter baumannii* as the major cause of ventilator-associated pneumonia in intensive care unit patients at an infectious disease hospital in southern Vietnam. *J Med Microbiol* 2014; **63:** 1386–94.

29 Salomão MC, Freire MP, Boszczowski I, Raymundo SF, Guedes AR, Levin AS. Increased risk for carbapenem-resistant Enterobacteriaceae colonization in intensive care units after hospitalization in emergency department. *Emerg Infect Dis* 2020; **26:** 1156–63.

30 Ingle DJ, Levine MM, Kotloff KL, Holt KE, Robins-Browne RM. Dynamics of antimicrobial resistance in intestinal *Escherichia coli* from children in community settings in south Asia and sub-Saharan Africa. *Nat Microbiol* 2018; **3:** 1063–73.