

1 **Title:** A genomic epidemiology study of multidrug-resistant *Escherichia coli*, *Klebsiella pneumoniae*  
2 and *Acinetobacter baumannii* in two intensive care units in Hanoi, Vietnam

3  
4 **Authors:**

5 Leah W. Roberts<sup>1,2</sup>, Le Thi Hoi<sup>3,4</sup>, Fahad A. Khokhar<sup>2,5</sup>, Nguyen Thi Hoa<sup>3,6</sup>, Tran Van Giang<sup>3,4</sup>, Cuong  
6 Bui<sup>7</sup>, Tran Hai Ninh<sup>3</sup>, Dao Xuan Co<sup>7</sup>, Nguyen Gia Binh<sup>7</sup>, Hoang Bao Long<sup>4</sup>, Dang Thi Huong<sup>3,4</sup>, James E.  
7 Bryan<sup>2</sup>, Archie Herrick<sup>2</sup>, Theresa Feltwell<sup>2</sup>, Behzad Nadjm<sup>8,9</sup>, H. Rogier van Doorn<sup>9,10</sup>, Julian Parkhill<sup>11</sup>,  
8 Nguyen Vu Trung<sup>3,4</sup>, Nguyen Van Kinh<sup>3</sup>, Zamin Iqbal<sup>1,\*</sup>, M. Estée Török<sup>2,12,\*</sup>

9 \*Contributed equally

10

11 **Affiliations:**

- 12 1. EMBL-EBI, Wellcome Genome Campus, Hinxton, Cambridge, United Kingdom
- 13 2. Department of Medicine, University of Cambridge, Cambridge, United Kingdom
- 14 3. National Hospital for Tropical Diseases, Hanoi, Vietnam
- 15 4. Hanoi Medical University, Hanoi, Vietnam
- 16 5. Cambridge Institute of Therapeutic Immunology and Infectious Diseases, University of  
17 Cambridge, Cambridge, United Kingdom
- 18 6. Department of Microbiology and National Tuberculosis Reference Laboratory, National Lung  
19 Hospital, Ha Noi, Viet Nam
- 20 7. Bach Mai Hospital, Hanoi Vietnam
- 21 8. Medical Research Council The Gambia at the London School of Hygiene & Tropical Medicine,  
22 Fajara, The Gambia
- 23 9. Wellcome Trust Major Overseas Programme, Hanoi, Vietnam
- 24 10. Nuffield Department of Medicine, University of Oxford, Oxford, UK
- 25 11. Department of Veterinary Medicine, University of Cambridge, Cambridge, United Kingdom
- 26 12. Departments of Infectious Diseases and Microbiology, Cambridge University Hospitals NHS  
27 Foundation Trust, Cambridge, United Kingdom

28

29 **Correspondence:**

30 M. Estée Török email: [et317@cam.ac.uk](mailto:et317@cam.ac.uk)

31 Zamin Iqbal, email: [zi@ebi.ac.uk](mailto:zi@ebi.ac.uk)

32

33 **Key words:** Antimicrobial resistance; multidrug-resistant; whole genome sequencing; transmission;

34 *Escherichia coli*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*; intensive care unit; Vietnam

**NOTE:** This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

35 **Abstract**

36 Background

37 Vietnam has high rates of antimicrobial resistance (AMR) but limited capacity for genomic  
38 surveillance. This study used whole genome sequencing (WGS) to examine the prevalence and  
39 transmission of three key AMR pathogens in two intensive care units in Hanoi, Vietnam.

40 Methods

41 A prospective surveillance study of all adults admitted to intensive care units (ICUs) at the National  
42 Hospital for Tropical Diseases (NHTD) and Bach Mai Hospital (BMH) was conducted between June  
43 2017 and January 2018. Clinical and environmental samples were cultured on selective media,  
44 characterised using MALDI TOF MS, and illumina sequenced. Phylogenies based on the *de novo*  
45 assemblies (SPAdes) were constructed using Mafft (PARsnp), Gubbins and RAxML. Resistance genes  
46 were detected using Abricate against the NCBI database.

47 Findings

48 3,153 *Escherichia coli*, *Klebsiella pneumoniae* and *Acinetobacter baumannii* isolates from 369 patients  
49 were analysed. Phylogenetic analysis revealed predominant lineages within *A. baumannii* (global  
50 clone [GC]2, sequence types [ST]2, ST571) and *K. pneumoniae* (ST15, ST16, ST656, ST11, ST147)  
51 isolates. Colonisation was most common with *E. coli* (88.9%) followed by *K. pneumoniae* (62.4%). Of  
52 the *E. coli*, 91% carried a *bla*CTX-M variant, while 81% of *K. pneumoniae* isolates carried *bla*NDM (54%)  
53 and/or *bla*KPC (45%). Transmission analysis using single nucleotide polymorphisms (SNPs) identified  
54 167 clusters involving 251 (68%) patients, in some cases involving patients from both ICUs. There were  
55 no significant differences between the lineages or AMR genes recovered between the two ICUs.

56 Interpretation

57 This study represents the largest prospective surveillance study of key AMR pathogens in Vietnamese  
58 ICUs. Clusters of closely related isolates in patients across both ICUs suggests recent transmission prior  
59 to ICU admission in other healthcare settings or in the community.

60 Funding

61 This work was funded by the Medical Research Council Newton Fund, United Kingdom; the Ministry  
62 of Science and Technology, Vietnam; and the Wellcome Trust, United Kingdom.

63

64

65

66

67

68

69

70

71

72 **Research in context**

73 Evidence before this study

74 Globally, antimicrobial resistance (AMR) is projected to cause 10 million deaths annually by 2050.  
75 Ninety percent of these deaths are expected to occur in low- and middle-income countries (LMICs),  
76 but attributing morbidity and mortality to AMR is difficult in the absence of comprehensive data.  
77 Whilst efforts have been made to improve AMR surveillance in these settings, this is often hampered  
78 by limited expertise, laboratory infrastructure and financial resources.

79

80 Added value of this study:

81 This is the largest prospective surveillance study of three key AMR pathogens (*E. coli*, *K. pneumoniae*  
82 and *A. baumannii*) conducted in critical care settings in Vietnam. Sampling was restricted to patients  
83 who were colonised or infected with extended spectrum beta-lactamase (ESBL) producing and/or  
84 carbapenem-resistant organisms. Colonisation with more than one organism was very common, with  
85 multidrug-resistant (MDR) *E. coli* being predominant in stool samples. A small number of predominant  
86 lineages were identified for *K. pneumoniae* and *A. baumannii*, while the *E. coli* isolates were highly  
87 genetically diverse. A large number of genomic clusters were identified within the two ICUs, some of  
88 which spanned both ICUs. There were no significant differences between lineages or AMR genes  
89 between the two ICUs.

90

91 Implications of all the available evidence

92 This study found high rates of colonisation and infection with three key AMR pathogens in adults  
93 admitted to two Vietnamese ICUs. Whilst transmission was common within ICUs the finding of similar  
94 lineages and AMR genes in both ICUs suggests that dissemination of AMR occurs prior to ICU  
95 admission, either in referral hospitals or in community settings prior to hospital admission. Strategies  
96 to tackle AMR in Vietnam will need to account for this by extending surveillance more widely across  
97 hospital and community settings.

98

99 **Introduction**

100 Low- and middle-income countries (LMICs) have reported widespread antimicrobial resistance (AMR)  
101 in healthcare, community and agricultural settings. In South-East Asia, dense human populations,  
102 intensive animal farming, unrestricted access to antibiotics and limited laboratory infrastructure have  
103 all contributed to the rapid expansion of AMR (1, 2).

104  
105 Much of this burden arises from excessive use of antimicrobials in human and animal populations. In  
106 Vietnam, antimicrobial usage has been estimated to be two times higher in humans, and 1.5 times  
107 higher in animals, as compared to the European Union (3). Despite legal restrictions in Vietnam,  
108 antibiotics are often dispensed without prescriptions in the community (4). Broad-spectrum  
109 antibiotics are also commonly administered in healthcare settings to mitigate the effects of limited  
110 capacity for microbiological testing and infection control (4, 5). Detection of both resistant bacteria  
111 and antimicrobials have been recorded in the environment (6, 7), hospital waste (8) and food sources  
112 (9, 10).

113  
114 Extensive AMR has led to increased pressure on hospitals and is particularly problematic in critical  
115 care settings. Whilst AMR surveillance based on phenotypic testing in Vietnam has improved in recent  
116 years, the infrastructure required for systematic genomic surveillance remains to be established.  
117 Genomic analysis is important to determine circulating lineages and elucidate potential transmission  
118 events, as other methods do not provide the same level of resolution (11). Over the past decade,  
119 studies have demonstrated the utility of WGS in characterising AMR, transmission routes, and  
120 dominant lineages (12-14). LMICs however remain relatively understudied, with few studies  
121 conducted in Vietnamese hospitals (15-18).

122  
123 In order to address this knowledge gap, we conducted a prospective genomic surveillance study of  
124 key AMR pathogens in two hospitals in Vietnam. We targeted intensive care unit (ICU) patients as we  
125 hypothesised that these would be most likely to have been treated with antibiotics and to harbour  
126 AMR pathogens. Furthermore, we focussed our analysis on the three most commonly isolated species  
127 (*Escherichia coli*, *K. pneumoniae*, and *A. baumannii*) that were extended-spectrum beta-lactamase  
128 (ESBL) producers and/or carbapenem-resistant. An additional in-depth analysis focussed primarily on  
129 a subset of the *K. pneumoniae* isolates was also performed in a separate project (Pham et al., personal  
130 communication).

131

132

## 133 **Methods**

### 134 *Study design, setting and participants*

135 This prospective observational cohort study was conducted in two hospitals, the National Hospital for  
136 Tropical Diseases (NHTD) and Bach Mai Hospital (BMH) in Hanoi, Vietnam, between June 2017 to  
137 January 2018. All adult patients admitted to the ICUs of the two hospitals were eligible for inclusion  
138 in the study. Screening specimens (stool/rectal swabs, urine, skin/wound swabs and sputum/tracheal  
139 aspirates) were collected from ICU patients on admission, on discharge and weekly during their ICU  
140 stay. Environmental samples were collected on a monthly basis.

141

### 142 *Laboratory methods*

143 Specimens were cultured on selective media to identify extended-spectrum beta-lactamase producers  
144 and carbapenem-resistant organisms. Target organisms (*Escherichia coli*, *Acinetobacter baumannii*  
145 and *Klebsiella pneumoniae*) were identified using MALDI-TOF MS and stored at -80 °C prior to being  
146 shipped to the University of Cambridge, United Kingdom, where they underwent antimicrobial  
147 susceptibility testing and DNA extraction.

148

### 149 *Sequencing and bioinformatic analysis*

150 DNA extracts were transferred to the Wellcome Sanger Institute for library preparation and  
151 sequencing. Further details of laboratory methods including read quality control, genome assembly,  
152 phylogenetic analysis, antibiotic resistance gene detection, multilocus sequence typing (MLST) and  
153 transmission cluster analysis are available in the Supplementary Materials.

154

## 155 **Results**

### 156 *Samples included in the study*

157 Between June 2017 to January 2018, a total of 3,367 isolates were cultured, comprising *Escherichia*  
158 *coli* (n=765), *Klebsiella pneumoniae* (n=1,372) and *Acinetobacter baumannii* (n=1,230). Thirty-one  
159 isolates were excluded from the analysis because of poor assembly quality. A further 150 isolates were  
160 excluded because of suspected inter-species contamination, and 33 isolates were excluded because  
161 of suspected intra-species (strain-level) contamination (Supplementary Figure 2). Thus 3,153 isolates  
162 (93.6%), comprising 2,901 isolates from 369 patients, and 252 environmental isolates, passed quality  
163 filtering and were included in the final analyses.

164

### 165 *Clinical characteristics*

166 3,153 bacterial isolates and 369 patients were included in the study. The number and type of samples  
167 and the participant baseline characteristics and outcomes are summarised in Table 1 and  
168 Supplementary Figure 3. The number of samples collected was higher, and the median length of ICU  
169 stay were longer at NHTD than BMH.

170

171 All of the 369 patients were colonised or infected with one or more of the three species (*E. coli*, *A.*  
172 *baumannii* and *K. pneumoniae*): 146 patients (40%, 55 at BMH and 91 at NHTD) with all three species;  
173 133 patients (36%, 66 at BMH and 67 at NHTD) with two of the three species; and 90 patients (24%,  
174 61 at BMH and 29 at NHTD) only had one species detected.

175

176 Both *E. coli* and *K. pneumoniae* were isolated primarily from stool/rectal swabs (627/721 [87.0%] and  
177 822/1316 [62.5%], respectively). *K. pneumoniae* was also isolated from other sites including sputum  
178 (325/1316 [24.7%]), urine samples (63/1316 [4.8%]) and skin swabs (17/1316 [1.3%]). In contrast, *A.*  
179 *baumannii* isolates were predominantly isolated from sputum (621/1116 [55.6%]), followed by  
180 stool/rectal swabs (247/1116 [22.1%]), urine (49/1116 [4.4%]) and skin swabs (36/1116 [3.2%]). *A.*  
181 *baumannii* also accounted for the highest number of environmental isolates (161/1116, [14.4%]),  
182 compared to 6.5% (85/1316) for *K. pneumoniae* and 2.2% (16/721) for *E. coli*.

183

#### 184 *Whole genome sequencing reveals dominant circulating lineages*

185 Phylogenetic trees for each species were constructed to explore lineage diversity. The *E. coli* isolates  
186 were found to be highly diverse, with isolates spread over eight phylogroups and 80 sequence types  
187 (STs) (Figure 1). The most prevalent ST was ST648 (phylogroup A; 11.8%), followed by ST410  
188 (phylogroup C; 9.7%), ST617 (phylogroup A; 9.2%), ST131 (phylogroup B2; 7.9%) and ST1193  
189 (phylogroup B2; 7.4%). Overall, 33 of the 80 STs only had one representative isolate in this dataset.

190

191 In contrast, the *K. pneumoniae* and *A. baumannii* isolates appeared to have a limited number of  
192 dominant lineages. More than 80% of the *K. pneumoniae* isolates were from one of five STs, including  
193 ST15 (n=34%), ST16 (n=20%), ST656 (n=12%), ST11 (n=11%) and ST147 (n=7%). The majority of *A.*  
194 *baumannii* were global clone 2 (GC2) (n=832, 74.6%) (19) and mainly belonged to ST2 (n=48%) and  
195 ST571 (n=24%) (based on the Pasteur scheme). There did not appear to be any specific relationship  
196 between STs and hospitals, with all of the major ST lineages represented in both ICUs.

197

198 In order to gain broader insight into the lineages, we selected globally representative strains to  
199 contextualise our dataset. Addition of these global representatives into the *E. coli* phylogeny showed

200 that most isolates belonged to a globally diverse set of STs that were not unique to Vietnam, but found  
201 across parts of North America, Europe and Asia (Supplementary Figure 4). Similarly, several of the  
202 major *K. pneumoniae* lineages were represented globally, particularly ST147, ST11 (mainly from China  
203 and the USA) and ST15 (mainly other Asian countries) (Supplementary Figure 5). However, it was also  
204 clear that local expansion was prominent, particularly among the ST656, ST16 and ST15 lineages. For  
205 *A. baumannii*, we focused primarily on GC2 isolates (Supplementary Figure 6). There was very little  
206 representation of global strains within our dataset, and those that were available consisted mainly of  
207 strains from other parts of Asia.

208  
209 Closer inspection of the global representatives found several strains in each species that were closely  
210 related (<5 core SNPs) to isolates in our dataset (Supplementary Table 1). Most of these global  
211 representatives were also isolated in Asian countries, with agriculture implicated as a source, based  
212 on two closely related *E. coli* representatives originally isolated from poultry (biosample  
213 SAMEA104188722) and a farm worker in Vietnam (biosample SAMEA5277968). There were only two  
214 closely related representatives isolated outside Asia, from the United Kingdom and Australia (*E. coli*  
215 ST1193 and ST131, respectively, which are both well-characterised global lineages).

216  
217 *High prevalence of antibiotic resistance genes among majority of isolates*

218 Almost all isolates carried acquired resistance genes belonging to at least three antibiotic classes, with  
219 90% of *E. coli*, 97% of *K. pneumoniae* and 41% *A. baumannii* carrying genes across five antibiotic classes  
220 (Figure 2). There were no discernible differences based on sample source or hospital, with the  
221 exception of *E. coli* detected in wound/skin swabs (n=6) which appeared on average to carry resistance  
222 to more antibiotic classes. *K. pneumoniae* isolates tended to fall into one of three “peaks” (Figure 2).  
223 This was due to lineage-specific carriage of acquired resistance genes, where ST15 isolates tended to  
224 carry resistance to more classes, compared to ST16 which often carried the least. The other three  
225 main lineages (ST656, ST11, ST147) fell between these two peaks. The exception was in the  
226 environmental samples, where only two peaks were seen and was likely due to very few ST16 isolates  
227 being detected in the environment.

228  
229 Resistance to antibiotics classes varied across the *E. coli* phylogeny, reflective of the diversity of strains  
230 within the dataset (Supplementary Figure 7). *bla*CTX-M genes were found in most *E. coli*, with *bla*CTX-  
231 M-15 (36%), *bla*CTX-M-27 (30%) and *bla*CTX-M-55 (17%) the most prevalent (Table 2). *bla*KPC-2 (13%)  
232 and *bla*NDM-[1,4,5,7] (24%) were present sporadically across the phylogroups, suggesting  
233 independent acquisitions events. Only 4% (n=28) of isolates carried *mcr* genes conferring resistance

234 to colistin. Again, these seemed to be independent acquisitions, with the exception of an ST206 cluster  
235 (phylogroup A; n=11) involving three patients from NHTD.

236

237 Conversely, MDR gene presence across the *K. pneumoniae* isolates appeared consistent with the main  
238 lineages, suggesting clonal expansion rather than diverse sampling of the species (Supplementary  
239 figure 8). Similar to the *E. coli*, incidence of *bla*CTX-M-15 (37.5%) was high, but less so than *bla*KPC-2  
240 (45%) and *bla*NDM (54%) (NDM-4 [27.9%], NDM-1 [24.7%] and NDM-5 [1.8%]) (Table 2).

241

242 Acquired AMR genes were overall less prevalent among the *A. baumannii* isolates. Similar to the *K.*  
243 *pneumoniae*, resistance to specific classes tended to be a feature of each distinct lineage, suggesting  
244 clonal expansion (Supplementary Figure 9). The carbapenemase gene *bla*OXA-23 was present in 83%  
245 of the dataset, with *bla*OXA-58 and *bla*OXA-72 present at much lower frequencies (5% and 0.2%  
246 respectively) (Table 2). The aminoglycoside resistance gene *armA* was also highly prevalent (76%).

247

248 Overall, 133 AMR genes were detected in BM and 154 were detected in NHTD. 49 genes were unique  
249 to either hospital (35 in NHTD, 14 in BM), but were only detected at a prevalence of less than 0.1%,  
250 suggesting sporadic cases. The remaining 129 genes were the same across both hospitals. Genes with  
251 at least 1% prevalence in each hospital were found to be almost identical. There were only five  
252 exceptions where the gene prevalence was >1% in BM but <1% in NHTD (*bla*NDM-4 (0.98%), *dfrA12*  
253 (0.92%), *rmtB1* (0.86%), *qnrB6* (0.74%) and *bla*OXA-181 (0.56%)). We also attempted to discover  
254 possible mobile genetic element (MGE)-driven AMR transmission by monitoring fluctuation of AMR  
255 genes prevalence over time (further explained in the supplementary results, supplementary figure  
256 10).

257

258 *Most patients carried several AMR strains*

259 For patients already infected or colonised with *E. coli*, *K. pneumoniae* and/or *A. baumannii*, over half  
260 had repeat isolation of the same bacterial species at a later timepoint during their stay (*E. coli*: 138/270  
261 [51%], *K. pneumoniae*: 169/294 [57%], *A. baumannii*: 133/230 [58%]). Of these patients, 60 to 70%  
262 had different sequence types (ST) (*E. coli*: 67%, *K. pneumoniae*: 68%, *A. baumannii*: 64%) (Figure 3).  
263 For *E. coli* and *K. pneumoniae*, the majority of patients only had isolates detected in stool (80% and  
264 54%, respectively). Conversely, most patients with *A. baumannii* were detected only in sputum (38%),  
265 or sputum and stool (32%).

266

267 *Evidence of extensive transmission between ICU patients*



268 Temporal observation of the isolates found no obvious association of any time point with any ST to  
269 suggest an outbreak of a specific lineage. In order to investigate potential clusters within the ICUs at  
270 a higher resolution, we examined plausible short-term transmission events using single nucleotide  
271 polymorphisms (SNPs).

272

273 To identify closely related strains that could indicate recent transmission, we evaluated clusters based  
274 on SNP distances across the core genome of each species for this dataset. Given the short sampling  
275 period, none of the three major species were likely to acquire more than 1 SNP while in the hospital.  
276 As such, we looked at samples with genomic evidence of most recent transmission; zero SNP clusters  
277 (follow-up with higher thresholds detailed below). Clusters were defined when they involved >1  
278 patient. Clusters involving a single patient and environmental samples were not included.

279

280 Most clusters were detected in *K. pneumoniae* and *A. baumannii* isolates, with 71 and 74 clusters  
281 representing 38% and 52% of total isolates for that species, respectively (Figure 4). *K. pneumoniae* had  
282 some of the largest clusters, ranging in size from 2 to 79 isolates, while *A. baumannii* clusters were  
283 smaller, between 2 to 33 isolates. Only 22 clusters were detected in *E. coli* and were generally small  
284 (median 3 isolates, range 2 to 9 isolates), representing only 13% of the *E. coli* dataset.

285

286 For all three species, the majority of clusters were detected between patients within a single ICU.  
287 Evaluating admission and discharge dates further confirmed patient overlap in these clusters (Figure  
288 4, Supplementary Figures 12 to 17). Additionally, between-ICU clusters were also detected (discussed  
289 further below).

290

291 Compared to *E. coli* and *K. pneumoniae*, *A. baumannii* clusters were more often associated with  
292 environmental isolates (24/74 clusters, Figure 4). The largest *A. baumannii* cluster within the same  
293 hospital ICU involved 33 isolates from eight patients and eight environmental samples (Supplementary  
294 Figure 13; ST451, cluster number 13). Admissions for patients in this cluster overlapped with detection  
295 of the same strain in the environment, which was also detected in the hospital environment several  
296 months later. Only two *E. coli* clusters contained related environmental isolates. *K. pneumoniae*  
297 environmental isolates were more often found in within-hospital (ICU) clusters (n=11) compared to  
298 between-hospital clusters (n=4).

299

300 To broadly evaluate infection risk for each outbreak cluster, we determine whether clusters contained  
301 (i) stool sample / rectal isolates only, indicating colonisation or (ii) isolates from urine, skin swab or

302 sputum samples, which could represent infection or colonisation at multiple sites. The majority of *A.*  
303 *baumannii* clusters contained isolates from non-stool / rectal swab samples (Figure 4). Conversely,  
304 colonisation-only clusters were common for *E. coli*; the largest *E. coli* cluster involving both hospital  
305 ICUs contained only 5 isolates from four patients, which were all isolated from stool (Supplementary  
306 Figure 14; ST617, cluster number 37). *K. pneumoniae* clusters were a mix, with both colonisation-only  
307 clusters and infection clusters.

308

309 In addition to suspected within-ICU transmission, we also detected a number of clusters involving  
310 patients from both hospital ICUs (Figure 4). The most pronounced example of this was a large ST15 *K.*  
311 *pneumoniae* cluster involving 79 isolates from 38 patients and 6 environmental samples  
312 (Supplementary Figure 18). Most were collected between July to September, with some late  
313 occurrences in October and November. All patients from NHTD between July to September had  
314 overlapping timelines, consistent with spread within the ICU (Supplementary figure 16; ST15, cluster  
315 number 15). Only one patient had no evidence of overlap (ND162) but did cluster with environmental  
316 isolates from the same timeframe, indicating a possible environmental source. Patients from BM  
317 appeared to have both patient overlap and consecutive acquisitions without patient overlap as time  
318 progressed. This suggests that patient-to-patient transmission and also transmission via other routes  
319 (e.g. inadequate cleaning before the next patient, transmission via healthcare workers) may have been  
320 important factors in the spread of this strain.

321

322 The identification of closely related isolates between independently operating ICUs suggested that  
323 there may have been a common source located outside the ICU e.g admission to the same location  
324 prior to admission to ICU. To determine whether certain lineages were more associated with  
325 acquisition within the ICU, we assessed diversity on arrival (i.e. the patients' first sample) versus  
326 diversity within the ICU (all other samples). Based on ST alone, we found a slight increase in diversity  
327 in the ICU versus on arrival (Supplementary Figure 19). However, the unique STs recovered in either  
328 setting only represented a small portion of the isolates overall. All of the main lineages for each species  
329 were found on both admission and within the ICU (Supplementary Figure 19 and 20).

330

### 331 *Identification of multiple transmission clusters per patients*

332 Overall, there were 251 patients (representing 68% of the cohort) involved in 167 clusters across the  
333 three species collected over the course of this study. 112 patients were only involved in a single cluster  
334 during their time in the ICU (Figure 5). However, the remaining 139 patients were involved in at least  
335 two clusters, with one patient involved in 12 clusters. For patients with at least two clusters, 20 had

336 clusters from all three species, 94 had clusters from two species and 25 had only one species. Overall,  
337 we saw a general trend towards more clusters in a single patient as they spent more time in the ICU  
338 ward.

339

340 To determine if any of our (zero SNP) clusters were potentially derived from a single original cluster  
341 predating their time in the ICU, we looked at SNP distances between clusters of the same ST  
342 (Supplementary Figure 21). At a threshold of five SNPs, several of the prominent STs within each  
343 species formed large clusters, including ST804 in *A. baumannii* and ST16 in *K. pneumoniae*. At this  
344 threshold, we found 29 clusters in the *A. baumannii* dataset (originally 74), 23 clusters in the *K.*  
345 *pneumoniae* (originally 71) and 19 clusters in *E. coli* (originally 22). By readjusting our analysis per  
346 patient using these clusters, we found that slightly more patients (123 versus 112 previously) may  
347 have only been involved in a single cluster during their stay (Supplementary figure 22). 128 patients  
348 had 2 clusters, with the maximum number of clusters in a single patient being 7 (n=3 patients).

349

## 350 **Discussion**

351

352 Here we present a large prospective genomic surveillance study of ESBL- or carbapenemase-producing  
353 organisms from three key AMR pathogens from two hospital ICUs in Vietnam. We used WGS to  
354 capture a high-resolution snapshot of the dominant circulating lineages over a six-month period. In  
355 this study we focused on *E. coli*, *K. pneumoniae*, and *A. baumannii*, as these were the most commonly  
356 isolated species from both ICUs. These three species have also been reported as highly prevalent in  
357 other Vietnamese studies (20), and are amongst the most clinically significant Gram-negative bacteria,  
358 having been designated as “critical” priority pathogens for research and development of new  
359 antibiotics by the World Health Organisation (21).

360

361 Despite limiting our study samples to ESBL-producing and/or carbapenem-resistant isolates belonging  
362 to three species, we identified a large number of isolates (an average of 17 and 10 isolates/day from  
363 patients in BMH and NHTD respectively). This is exceedingly high compared to other countries, such  
364 as the United Kingdom (UK), where a similar study only found 199 ESBL-producing Enterobacteriaceae  
365 over the course of one year (0.5 isolates/day) from three hospital sites between 2008 to 2009 (22). A  
366 point-prevalence survey conducted in a UK hospital in 2017 also identified no positive  
367 carbapenemase-producing Enterobacteriaceae (CPE) from 540 samples (23).

368

369 Colonisation was found to be a large reservoir for AMR, with the majority of *E. coli* and more than half  
370 of the *K. pneumoniae* isolated from stool samples, as has been documented previously in Vietnamese  
371 hospitals (24). High community usage of antibiotics in Vietnam is likely to promote colonisation with  
372 AMR bacteria, as prior treatment with antibiotics is known to lead to colonisation (25). Colonisation  
373 itself has been identified as a risk factor for subsequent infection (26-28), and has been previously  
374 documented in ICU patients in southern Vietnam (29). Colonisation with AMR *E. coli* has also been  
375 identified as a risk for transferral of resistance to other colonising pathogens, such as *Shigella* (30).

376

377 In addition to evidence for recent transmission between patients on the same ICU, we also identified  
378 clusters involving patients from both hospital ICUs. This result was unexpected, as there was no direct  
379 transfer of patients between the two ICUs. The most likely explanation for these clusters is a source  
380 outside of the ICUs. This might be other wards or hospitals which may have referred patients to  
381 intensive care, or AMR strains that have been acquired in the community, reflecting high rates of  
382 antibiotic use. Based on the similarity between lineages and AMR across both ICUs, we suggest that  
383 transmission, particularly of the predominant *A. baumannii* and *K. pneumoniae* lineages, is likely  
384 already circulating outside of the ICUs, where it is then further propagated.

385

386 In this context, we should potentially reconsider how we model AMR surveillance and control.  
387 Knowing whether hospitals are the primary source of AMR bacteria (and subsequent transmission into  
388 the community) (31), or whether the high rates are part of a more general, endemic pattern of  
389 circulating resistant strains is important for our understanding of local transmission dynamics. If the  
390 latter is true, this leads to very different national and indeed global management plans.

391

392 We acknowledge some limitations to our study. Firstly, we did not extensively explore plasmid profiles  
393 amongst the samples because of the limitations of short read sequence data. Secondly, although serial  
394 samples were collected and cultured, we selected single colony picks for sequencing. It is therefore  
395 possible that some diversity may have been lost at different timepoints during the study sampling  
396 period. Finally, this study focused on patient and environmental samples only. We were therefore  
397 unable to investigate potential transmission events involving hospital staff and/or visitors.

398

399 Nevertheless, we present the largest prospective surveillance study of multidrug-resistant *E. coli*, *A.*  
400 *baumannii* and *K. pneumoniae* in Vietnamese critical care patients to date. The extensive transmission  
401 and AMR detected within and between ICU wards suggests dominant circulating lineages of *A.*  
402 *baumannii* and *K. pneumoniae* existing both within hospitals, and potentially in community settings in

403 Vietnam. Further work is required to expand genomic surveillance in hospital and community settings  
404 in order to inform AMR control strategies in Vietnam.

405

406 *Data sharing*

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

411

412 *Ethical statement*

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

417

418 *Declaration of interests*

419 All authors declare no conflicts of interest

420

421 *Funding statement*

422 This work was funded by the Medical Research Council Newton Fund, the Vietnamese Ministry of  
423 Science and Technology (HNQT/SPĐP/04.16), and the Wellcome Trust. LWR is supported by an EMBL-  
424 EBI biomedical postdoctoral research fellowship. MET was supported by a Clinician Scientist  
425 Fellowship (funded by the Academy of Medical Sciences and the Health Foundation) and the NIHR  
426 Cambridge Biomedical Research Centre. The funding source had no role in study design; in the  
427 collection, analysis, and interpretation of data; in the writing of the report; and in the decision to  
428 submit the paper for publication.

429

430 **Author contributions**

431 Conceptualisation: LWR, JP, NVK, ZI, MET

432 Data collection: FK, NTHM, LTH, NGB, DXC, LTH, NTH, TG, CB, HNT, BN, RvD, NVT

433 Sample processing: NTH, FK, JB, JH, TF

434 Methodology: LWR, ZI

435 Formal Analysis: LWR, ZI

436 Writing (original draft): LWR

437 Writing (review/editing): ZI, MET  
438 Supervision: NVT, RvD, NVK, ZI, MET  
439 Project administration: LTH, NVT, RvD, MET  
440 Funding acquisition: MET, NVK and JP

441

#### 442 **Acknowledgements**

443 The authors would like to thank the patients for participating in this study and the clinical and  
444 laboratory staff of the National Hospital for Tropical Diseases and Bach Mai Hospital for their  
445 assistance with this study. We also acknowledge the sequencing team at the Wellcome Sanger  
446 Institute for their assistance with sequencing the samples including in the study.

447

#### 448 **References**

- 449 1. R. M. Zellweger *et al.*, A current perspective on antimicrobial resistance in Southeast  
450 Asia. *J Antimicrob Chemother* **72**, 2963-2972 (2017).
- 451 2. B. A. Walther *et al.*, Biodiversity and health: Lessons and recommendations from an  
452 interdisciplinary conference to advise Southeast Asian research, society and policy.  
453 *Infect Genet Evol* **40**, 29-46 (2016).
- 454 3. J. J. Carrique-Mas, M. Choisy, N. Van Cuong, G. Thwaites, S. Baker, An estimation of  
455 total antimicrobial usage in humans and animals in Vietnam. *Antimicrob Resist Infect*  
456 *Control* **9**, 16 (2020).
- 457 4. K. V. Nguyen *et al.*, Antibiotic use and resistance in emerging economies: a situation  
458 analysis for Viet Nam. *BMC Public Health* **13**, 1158 (2013).
- 459 5. M. E. de Kraker, A. J. Stewardson, S. Harbarth, Will 10 Million People Die a Year due  
460 to Antimicrobial Resistance by 2050? *PLoS Med* **13**, e1002184 (2016).
- 461 6. N. H. Tran *et al.*, Occurrence and risk assessment of multiple classes of antibiotics in  
462 urban canals and lakes in Hanoi, Vietnam. *Sci Total Environ* **692**, 157-174 (2019).
- 463 7. S. Yamasaki, T. D. Le, M. Q. Vien, C. Van Dang, Y. Yamamoto, Prevalence of  
464 extended-spectrum beta-lactamase-producing *Escherichia coli* and residual  
465 antimicrobials in the environment in Vietnam. *Anim Health Res Rev* **18**, 128-135  
466 (2017).
- 467 8. H. A. Duong *et al.*, Occurrence, fate and antibiotic resistance of fluoroquinolone  
468 antibacterials in hospital wastewaters in Hanoi, Vietnam. *Chemosphere* **72**, 968-973  
469 (2008).
- 470 9. N. T. Nhung *et al.*, Antimicrobial residues and resistance against critically important  
471 antimicrobials in non-typhoidal *Salmonella* from meat sold at wet markets and  
472 supermarkets in Vietnam. *Int J Food Microbiol* **266**, 301-309 (2018).
- 473 10. S. P. T. S. S. A. K. Anal, Addressing the antibiotic resistance and improving the food  
474 safety in food supply chain (farm-to-fork) in Southeast Asia. *Food Control* **108**,  
475 (2020).
- 476 11. S. Quainoo *et al.*, Whole-Genome Sequencing of Bacterial Pathogens: the Future of  
477 Nosocomial Outbreak Analysis. *Clin Microbiol Rev* **30**, 1015-1063 (2017).

- 478 12. M. S. Toleman *et al.*, Prospective genomic surveillance of methicillin-resistant  
479 Staphylococcus aureus (MRSA) associated with bloodstream infection, England, 1  
480 October 2012 to 30 September 2013. *Euro Surveill* **24**, (2019).
- 481 13. H. T. T. Phan *et al.*, Illumina short-read and MinION long-read WGS to characterize  
482 the molecular epidemiology of an NDM-1 *Serratia marcescens* outbreak in Romania.  
483 *J Antimicrob Chemother* **73**, 672-679 (2018).
- 484 14. E. M. Harrison *et al.*, Transmission of methicillin-resistant Staphylococcus aureus in  
485 long-term care facilities and their related healthcare networks. *Genome Med* **8**, 102  
486 (2016).
- 487 15. A. Santana *et al.*, Emergence of unusual vanA/vanB2 genotype in a highly mutated  
488 vanB2-vancomycin-resistant hospital-associated *E. faecium* background in Vietnam.  
489 *Int J Antimicrob Agents* **52**, 586-592 (2018).
- 490 16. H. Chung The *et al.*, A high-resolution genomic analysis of multidrug-resistant  
491 hospital outbreaks of *Klebsiella pneumoniae*. *EMBO Mol Med* **7**, 227-239 (2015).
- 492 17. B. Berglund *et al.*, Colistin- and carbapenem-resistant *Klebsiella pneumoniae* carrying  
493 mcr-1 and blaOXA-48 isolated at a paediatric hospital in Vietnam. *J Antimicrob*  
494 *Chemother* **73**, 1100-1102 (2018).
- 495 18. B. Berglund *et al.*, Molecular and phenotypic characterization of clinical isolates  
496 belonging to a KPC-2-producing strain of ST15 *Klebsiella pneumoniae* from a  
497 Vietnamese pediatric hospital. *Antimicrob Resist Infect Control* **8**, 156 (2019).
- 498 19. J. F. Turton *et al.*, Identification of *Acinetobacter baumannii* by detection of the  
499 blaOXA-51-like carbapenemase gene intrinsic to this species. *J Clin Microbiol* **44**,  
500 2974-2976 (2006).
- 501 20. T. V. D. Vu *et al.*, Antimicrobial susceptibility testing and antibiotic consumption  
502 results from 16 hospitals in Viet Nam: The VINARES project 2012-2013. *J Glob*  
503 *Antimicrob Resist* **18**, 269-278 (2019).
- 504 21. W. H. Organisation. (2017), vol. 2020.
- 505 22. D. A. Enoch *et al.*, Epidemiology of extended-spectrum beta-lactamase-producing  
506 Enterobacteriaceae in a UK district hospital; an observational study. *J Hosp Infect* **81**,  
507 270-277 (2012).
- 508 23. H. J. Wilson *et al.*, Point-prevalence survey of carbapenemase-producing  
509 Enterobacteriaceae and vancomycin-resistant enterococci in adult inpatients in a  
510 university teaching hospital in the UK. *J Hosp Infect* **100**, 35-39 (2018).
- 511 24. D. M. Tran *et al.*, High prevalence of colonisation with carbapenem-resistant  
512 Enterobacteriaceae among patients admitted to Vietnamese hospitals: Risk factors  
513 and burden of disease. *J Infect* **79**, 115-122 (2019).
- 514 25. J. Ruiz *et al.*, Influence of antibiotic pressure on multi-drug resistant *Klebsiella*  
515 *pneumoniae* colonisation in critically ill patients. *Antimicrob Resist Infect Control* **8**,  
516 38 (2019).
- 517 26. C. M. Manaia, Assessing the Risk of Antibiotic Resistance Transmission from the  
518 Environment to Humans: Non-Direct Proportionality between Abundance and Risk.  
519 *Trends Microbiol* **25**, 173-181 (2017).
- 520 27. M. E. Flokas, S. A. Karageorgos, M. Detsis, M. Alevizakos, E. Mylonakis, Vancomycin-  
521 resistant enterococci colonisation, risk factors and risk for infection among  
522 hospitalised paediatric patients: a systematic review and meta-analysis. *Int J*  
523 *Antimicrob Agents* **49**, 565-572 (2017).

- 524 28. J. Asir *et al.*, Simultaneous gut colonisation and infection by ESBL-producing  
525 Escherichia coli in hospitalised patients. *Australas Med J* **8**, 200-207 (2015).
- 526 29. D. B. Thuy, Colonisation and Infection with Antibiotic-Resistant Organisms in Patients  
527 in the Adult Intensive Care Unit, Hospital for Tropical Diseases, Ho Chi Minh City,  
528 Vietnam. *PhD thesis, University of Oxford*, (2018).
- 529 30. P. Thanh Duy *et al.*, Commensal Escherichia coli are a reservoir for the transfer of  
530 XDR plasmids into epidemic fluoroquinolone-resistant Shigella sonnei. *Nat Microbiol*  
531 **5**, 256-264 (2020).
- 532 31. S. David *et al.*, Epidemic of carbapenem-resistant Klebsiella pneumoniae in Europe is  
533 driven by nosocomial spread. *Nat Microbiol* **4**, 1919-1929 (2019).
- 534
- 535



536 **Display items**

537 **Table 1: Summary of study samples and participants**

	Bach Mai Hospital	National Hospital of Tropical Diseases
<b>Samples</b>		
Total	1042	2111
Clinical	993 (95%)	1898 (90%)
Environmental	49 (5%)	213 (10%)
<b>Study participants</b>		
Number	182	187
Male sex	104 (57%)	114 (61%)
Female sex	71 (39%)	50 (27%)
Sex not recorded	7 (4%)	23 (12%)
<b>Age</b>		
Male	53 years (range 16-85)	55.2 years (range 5-92)
Female	53.4 years (range 19-91)	55.7 years (range 10-90)
Age not recorded	n=5 (3%)	n=19 (10%)
<b>Duration of ICU admission</b>		
Median length of stay (LOS)	7.5 days (range 0 – 35)	21 days (range 1 – 75)
LOS not recorded	n=4 (2.2%)	n=7 (3.7%)
<b>Outcome at ICU discharge</b>		
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%)

538

539

540

541

542 **Table 2. Summary of resistance genes found in the three species**

Resistance gene class	<i>E. coli</i>	<i>A. baumannii</i>	<i>K. pneumoniae</i>
Tetracycline	563 (78.1%)	701 (62.8%)	753 (57.2%)
Sulphonamide	649 (90%)	746 (66.8%)	969 (73.6%)
Fluoroquinolone	161 (22.3%)	19 (1.7%)	1187 <sup>3</sup> (90.2%)
Colistin	28 (3.9%)	0 (0%)	10 (0.8%)
Fosfomycin	48 (6.7%)	5 (0.4%)	1316 <sup>3</sup> (100%)
MLS	579 (80%)	816 (73.1%)	623 (47.3%)
Trimethoprim	621 (86.1%)	41 (3.7%)	982 (74.6%)
Phenicols	274 (38%)	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%)
<i>Class C:</i>			
EC	721 <sup>1</sup>	2	0
ACT	0	1	0
CMY	209	1	2
DHA	26	2	24
<i>Class A:</i>			
LAP	17	0	142
CARB	0	63	0
PER	0	68	0
TEM	347	697	686
SHV	7	9	1292 <sup>3</sup>
VEB	0	12	3
CTX	613	4	681
KPC	94	3	593
<i>Class D:</i>			
OXA	251	996 <sup>2</sup>	611
<i>Class B:</i>			
IMP	0	6	1
NDM	173	35	716
Rifamycin	114 (15.8%)	63 (5.6%)	810 (61.6%)
Aminoglycoside	680 (94.3%)	1115 (99.9%)	1290 (98%)
Streptothricin	11 (1.53%)	8 (0.7%)	0 (0%)

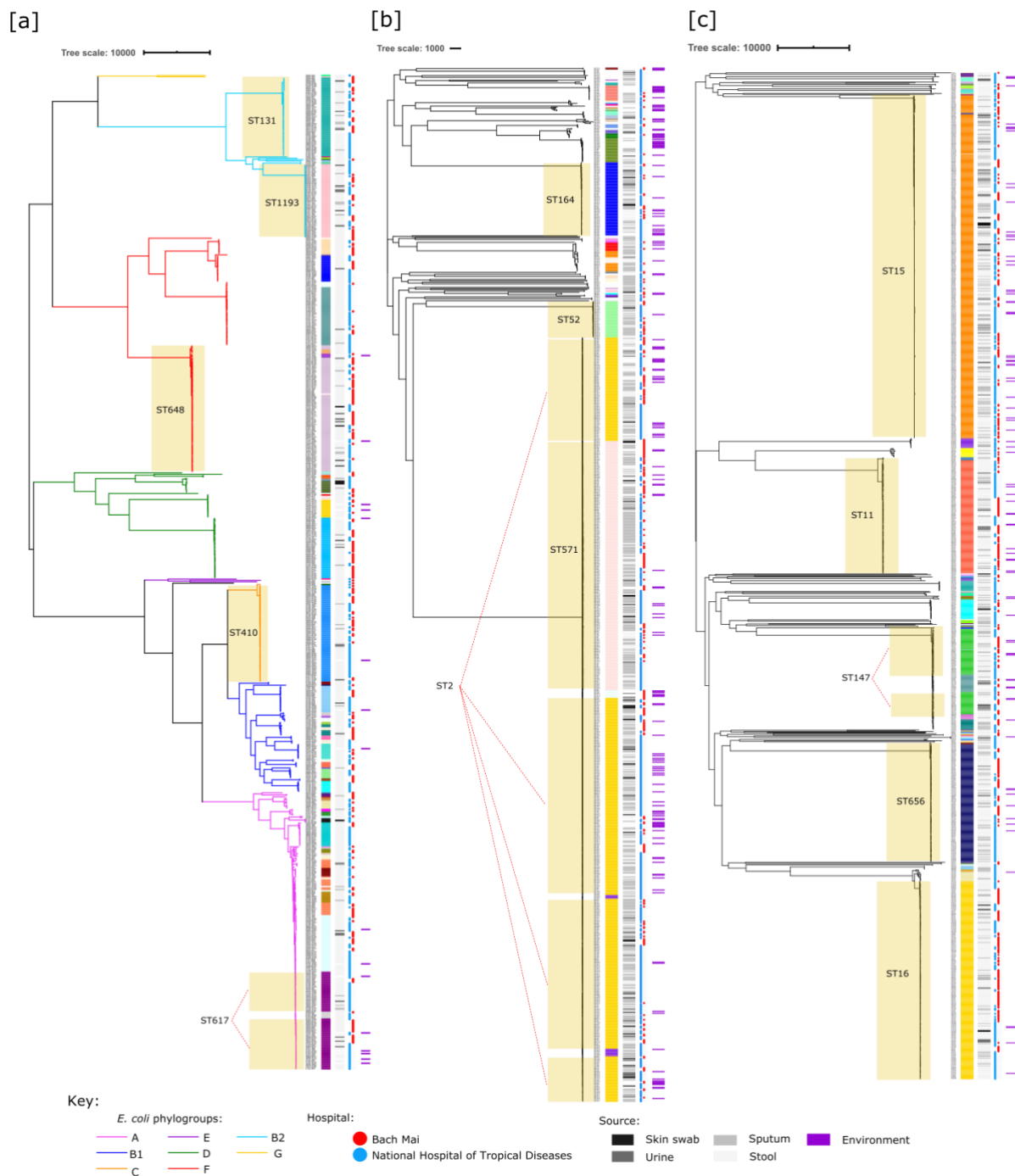
543 <sup>1</sup>blaEC intrinsic in *E. coli*

544 <sup>2</sup>blaADC and blaOXA intrinsic in *A. baumannii* (except OXA-[1,10,23,58,72])

545 <sup>3</sup>fosA (fosfomycin), oqxAB (fluoroquinolone) and blaSHV intrinsic in *K. pneumoniae*

546 ESBL resistance genes are highlighted in grey and carbapenem resistance genes in yellow

547 **Figure 1.**



548

549

550

551

552

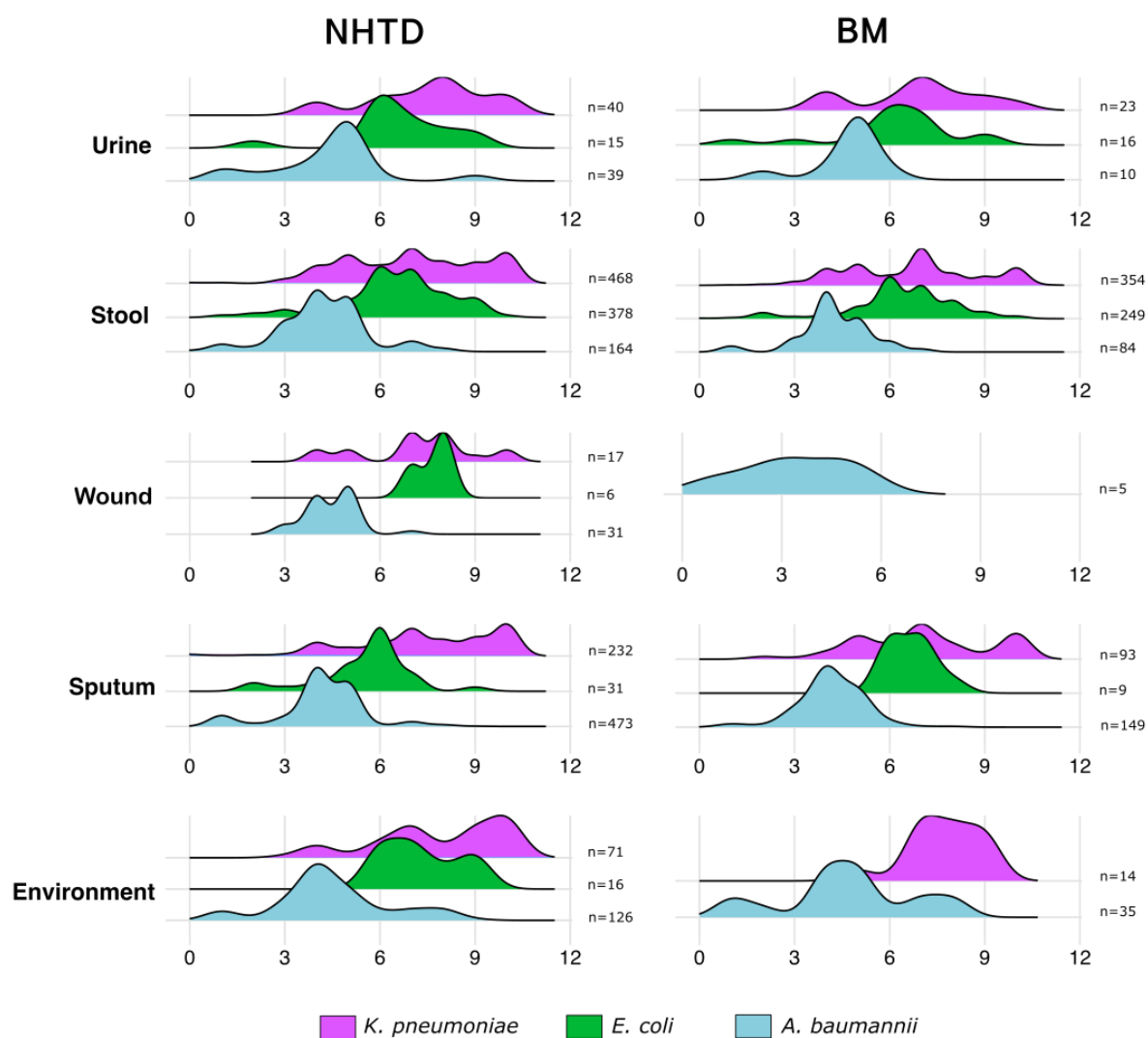
553

554

555

**Figure 1: Whole genome phylogenies for [A] *E. coli*, [B] *A. baumannii*, and [C] *K. pneumoniae*:** recombination-filtered core-SNP trees with mid-point root. Tree metadata includes (from left to right column beside trees): MLST, source and hospital. Outermost purple bars indicate environmental isolates. Branches corresponding to *E. coli* phylogroups are coloured accordingly. Main STs are highlighted in the image using the pale-yellow boxes.

556 **Figure 2.**



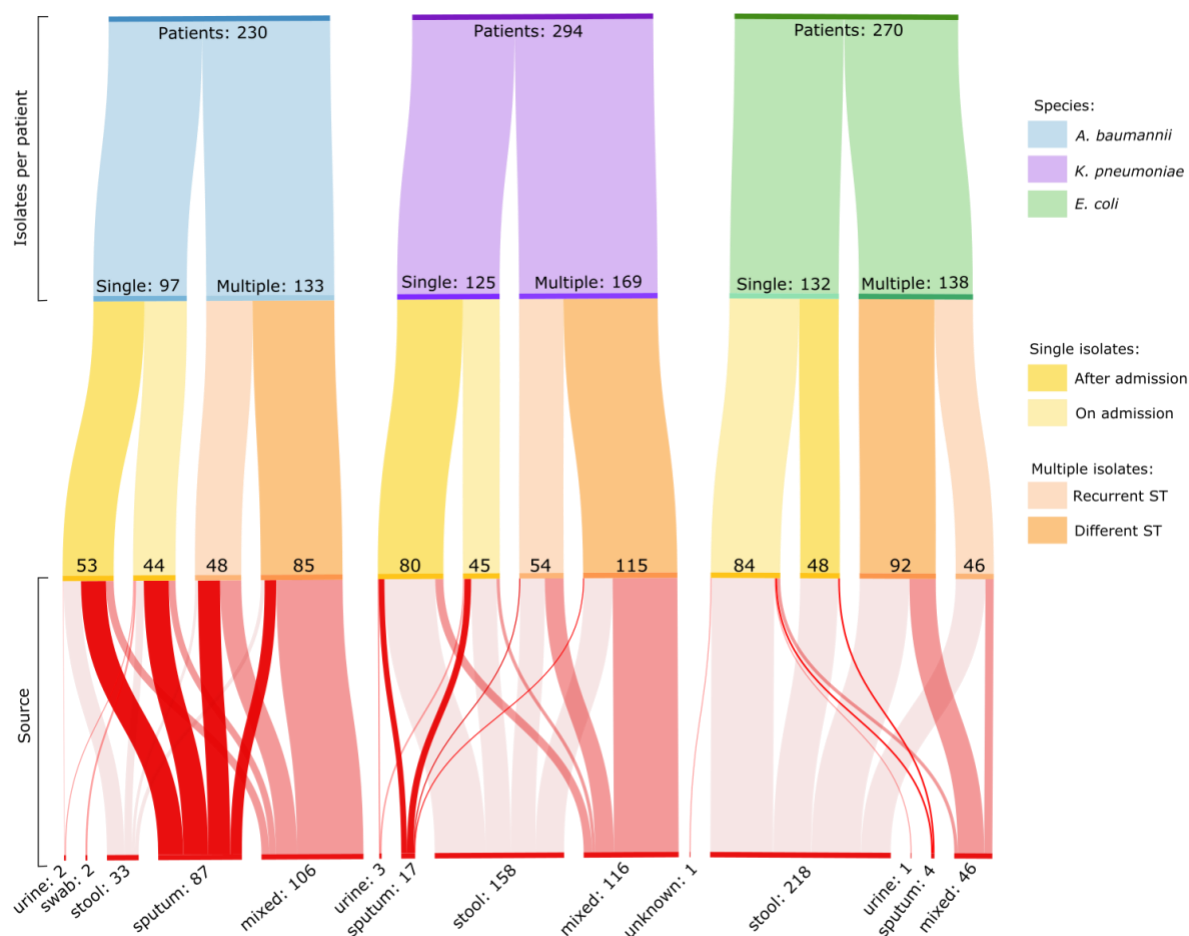
557

558 **Figure 2:** Summary of isolates and the number of antibiotic resistance classes separated by species, hospital and sample

559 type.

560

561 **Figure 3.**



562

563 **Figure 3: Overview of strain diversity, recurrence and source among study patients: “Patients”** refers to the total number

564 of patients in this study that had at least one isolate of that species. Within each species, we evaluated whether patients had

565 (i) only a single isolate for that species, or (ii) multiple isolates. If only a single isolate, we determined whether it was collected

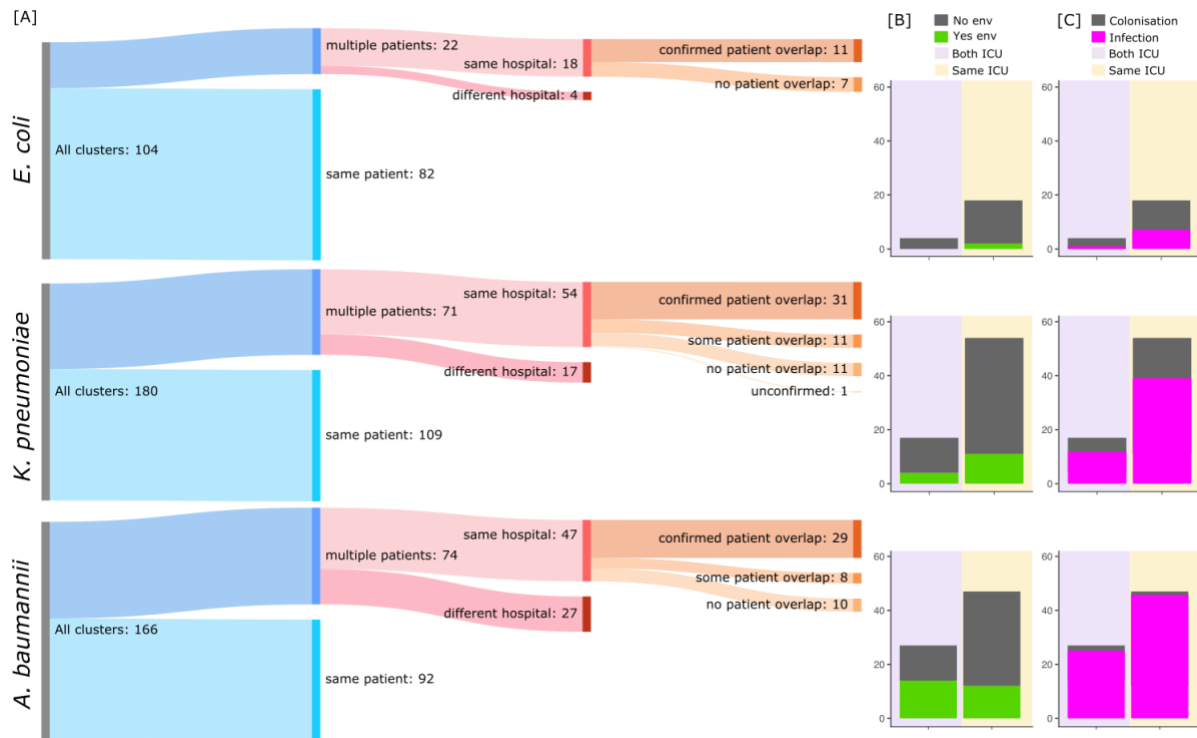
566 on admission to the ICU or after. For multiple isolates, we determined if the patient’s isolates were the same ST (recurrent)

567 or a different ST. We finally looked at how many patients had isolates from one site (urine, swab, stool or sputum) or mixed

568 sites (any combination of sites).

569

570 **Figure 4.**



571

572

573

574

575

576

577

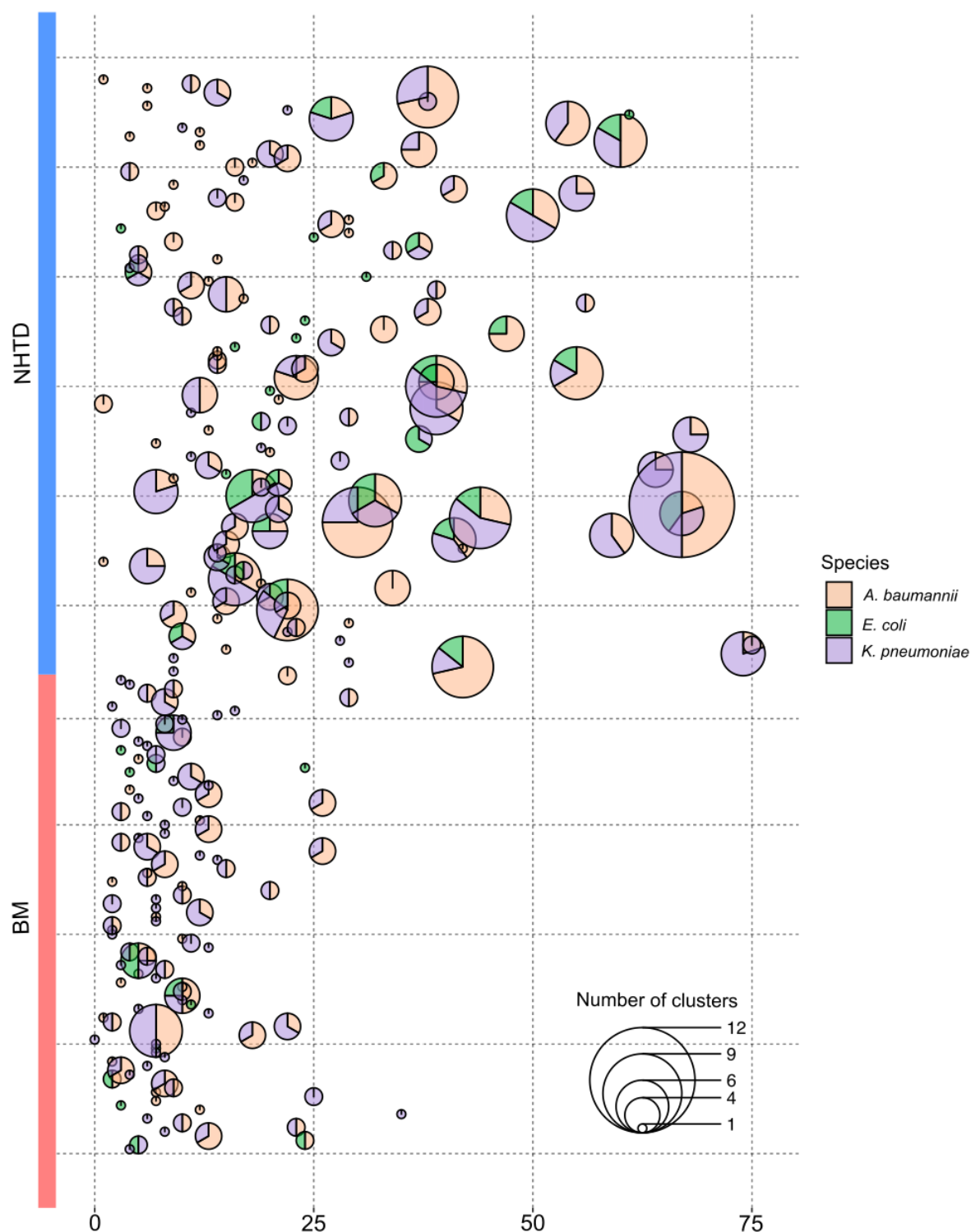
578

579

580

**Figure 4: Summary of 0 SNP clusters in all species:** [A] Clusters were defined as (i) multiple patients: samples were derived from at least two different patients, or (ii) 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 (i) confirmed patient overlap: all patient ICU stays overlap with another in the same cluster, (ii) some patient overlap: at least 2 patient ICU stays overlap, and (iii) zero patient overlap between all patients in cluster. [B] Environmental isolates in clusters: clusters were counted if an environmental isolate was found in that cluster. Y-axis = number of clusters. [C] Colonisation vs. infection: clusters were counted if they (a) had only isolates from stool (i.e. colonisation) or (b) had isolates from urine, swabs and/or sputum with or without isolates from stool (i.e. infection). Y-axis = number of clusters.

581 **Figure 5.**



582

583 **Figure 5: Scatterpie showing the number of clusters in patients across all species:** y-axis represents patients from BM or

584 NHTD involved in at least one 0 SNP cluster. X-axis represents length of stay for that patient; one pie is plotted per patient

585 at the duration of their stay. Each circle represents 0 SNP clusters in a single patient. The size of the clusters corresponds to

586 the number of clusters, while the colour corresponds to the species.

587