Genomic epidemiology and longitudinal sampling of ward wastewater environments and patients reveals complexity of the transmission dynamics of *bla*$_{KPC}$-carbapenemase-producing Enterobacterales in a hospital setting

Stoesser N[a, b, *], George R[c, *], Aiken Z[c], Phan HTT[d], Lipworth S[a], Wyllie DH[e], Quan TP[a], Mathers AJ[f], De Maio N[g], Seale AC[h], Eyre DW[a, b], Vaughan A[a], Swann J[a], Peto TEA[a, i], Crook DW[a, b, i], Cawthorne J[c], Dodgson A[c], Walker AS[a, b, i], and the TRACE Investigators' Group

[a] Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom

[b] NIHR Health Protection Research Unit in Healthcare Associated Infections and Antimicrobial Resistance at University of Oxford in partnership with Public Health England, Oxford, United Kingdom

[c] Manchester University NHS Foundation Trust, Manchester, United Kingdom

[d] University of Southampton, Southampton, United Kingdom

[e] United Kingdom Health Security Agency, United Kingdom

[f] University of Virginia, Charlottesville, United States of America

[g] EMBL-European Bioinformatics Institute, Cambridge, United Kingdom

[h] University of Warwick, Coventry, United Kingdom

[i] NIHR Oxford Biomedical Research Centre, Oxford University Hospitals NHS Foundation Trust, John Radcliffe Hospital, Oxford, United Kingdom

*Contributed equally

Keywords: Carbapenemase-producing Enterobacterales, genomic epidemiology, *bla*$_{KPC}$, wastewater sites, healthcare-associated

Corresponding author: Nicole Stoesser

Corresponding address: Department of Microbiology (Research), Level 7, John Radcliffe Hospital, Headley Way, Headington, OX3 9DU

Affiliation: Nuffield Department of Medicine, University of Oxford, Oxford UK and Oxford University Hospitals NHS Foundation Trust, Oxford, UK

e-mail: nicole.stoesser@ndm.ox.ac.uk

34 **ABSTRACT**

35 **Background**

36 Healthcare-associated wastewater reservoirs and asymptomatic gastrointestinal
37 patient colonisation by carbapenemase-producing Enterobacterales (CPE) can be
38 important in nosocomial CPE dissemination and infection. We characterised these
39 niches and within-niche diversity in a *bla*$_{KPC}$-associated CPE (KPC-E) endemic
40 healthcare setting, to better understand transmission potential.

41 **Methods**

42 We systematically sampled wastewater sites and patients across three units (six
43 wards) over 6-12 months in 2016 in a KPC-E endemic hospital. We used Illumina
44 sequencing to characterise up to five isolates per sample. Recombination-adjusted
45 phylogenies were used to define genetically related strains; assembly and mapping-
46 based typing approaches were used to characterise antimicrobial resistance gene and
47 insertion sequence profiles, and Tn*4401* types/target site sequences. The wider
48 accessory genome was evaluated in a subset of the largest clusters, and those crossing
49 niches.

50 **Findings**

51 Wastewater site KPC-E-positivity was substantial (101/349 sites [28.9%] positive,
52 319/4,488 [7.1%] sampling events positive); 183/4,425 (4.1%) of patients were CPE
53 culture-positive over the same timeframe. 13 species and 109 KPC-E strains were
54 observed across niches, and 24% of wastewater and 26% of patient KPC-E-positive
55 samples harboured ≥1 strain. Most diversity was explained by the individual niche,
56 suggesting highly localised factors are important in selection and spread.
57 Tn*4401*+target site sequence (TSS) diversity was greater in wastewater sites
58 ($p<0.001$), suggesting these might favour Tn*4401*-associated transposition/evolution

and dissemination. Shower/bath and sluice/mop-associated sites were more likely to be KPC-E-positive (Adjusted Odds Ratio [95% CI]: 2.69 [1.44-5.01], p=0.0019 and 2.60 [1.04-6.52], p=0.0410, respectively). Different strains had different transmission and bla_{KPC} dissemination dynamics.

Interpretation

There may be substantial KPC-E colonisation of wastewater sites and patients in KPC-E-endemic healthcare settings. Niche-specific factors, and different strains with different transmission dynamics influence carbapenemase gene dissemination. New transmission models incorporating complex, multi-level dynamics are needed to better quantify CPE dissemination to inform interventions and reduce transmission.

Funding

This study was supported by the National Institute for Health Research, UK.

71 **RESEARCH IN CONTEXT**

72 **Evidence before this study**

73 We searched PubMed to identify previous studies using whole genome sequencing
74 (WGS) to characterise in-hospital dissemination and persistence of carbapenemase-
75 producing Enterobacterales (CPEs) or the presence of CPEs in hospital wastewater
76 reservoirs. In our search (01/Jan/1996-30/Sep/2021) we used the terms ((((CPE) OR
77 (KPC)) AND (healthcare-associated OR nosocomial OR hospital-associated)) AND
78 (genom*) AND ((environment OR patient))). We had no restrictions on language. We
79 found sixty-six studies, but these were limited by: (i) evaluating single species or
80 lineages only (n=34), (ii) including patients or environmental outbreaks only, and not
81 both (n=23), (iii) sampling only small numbers (n<50) of patients or environmental
82 sites (n=26), or (iv) performing no evaluation of within-niche diversity (n=64). 13
83 studies were not relevant (e.g. investigating non-hospital locations, *ex vivo* or *in vivo*
84 evolution), two studies did not use WGS, and one study was not primary research.

85 **Added value of this study**

86 We provide a comprehensive assessment of healthcare-associated CPE dissemination
87 and persistence in hospital wastewater reservoirs through systematic sampling of a
88 large number of environmental sites (n=349 sites, n=4,488 samples) and patients
89 (n=4,425) over 6-12 months in a single hospital. For all CPE-positive environmental
90 samples (n=319) and a subset of CPE-positive patient samples (n=97/399 [24.3%]
91 samples from 76/183 [41.5%] CPE-positive patients), we sampled up to five CPE
92 isolates per sample to capture within-niche diversity, and used short-read Illumina
93 WGS (n=1,732 isolates successfully sequenced) to characterise within-niche
94 diversity, changes in colonisation over time, and genetic overlap between patient and
95 environmental niches.

96 **Implications of all the available evidence**

97 Through a detailed and resource-intensive study we captured the dynamics of seeding
98 and dissemination of important carbapenemase genes amongst patients and
99 environmental reservoirs within a hospital. We found differential colonisation and
100 dissemination dynamics for different species and lineages within and between niches.
101 Improved approaches incorporating variable within-niche diversity, accessory
102 distances and horizontal gene transfer in transmission evaluations are required to
103 better understand CPE dissemination and persistence, in order to direct interventions
104 limiting transmission.

105 **INTRODUCTION**

106 Carbapenemase-producing Enterobacterales (CPE) are a global health threat[1], and
107 commonly associated with multi-drug resistance, rendering treatment of CPE
108 infections difficult. Major transmissible global carbapenemases include the metallo-
109 beta-lactamases (bla_{NDM}, bla_{VIM}, bla_{IMP}), some oxacillinases (blaOXA$_{-48/48\text{-like}}$
110 variants), and the *Klebsiella pneumoniae* carbapenemase (KPC, encoded by bla_{KPC})[2].
111 Intra- and inter-species horizontal transfer of these genes is facilitated by mobile
112 genetic elements such as transposons and plasmids, resulting in rapid dissemination of
113 carbapenem resistance[3, 4].
114
115 CPE have been isolated from human and animal gastrointestinal tracts, and from
116 sewage, rivers and sink drains[5-7]. Recent studies have highlighted that hospital
117 wastewater sites may act as a CPE reservoir[5], but the diversity within these niches,
118 genetic overlap with patient isolates and likely directionality and rates of transmission
119 remain to be elucidated. Studies evaluating the contribution of patient-to-patient
120 transmission as the sole explanatory factor explaining transmission have generally
121 been unable to robustly explain the majority of transmission events, suggesting that
122 patient, staff and/or environmental reservoirs remain insufficiently sampled to
123 account for the diversity observed[8]. Few studies have considered within-niche
124 diversity of strains by sampling multiple isolates per individual and/or site[9-11].
125
126 The Manchester University NHS Foundation Trust (MFT; Manchester, UK), has
127 experienced on-going transmission and patient acquisition of multi-species CPE,
128 predominantly bla_{KPC}-positive Enterobacterales (KPC-E), since 2009. Following a
129 KPC-*Escherichia coli* outbreak on two cardiac wards in the Manchester Royal

130 Infirmary (MRI; part of MFT), and subsequent ward closure and plumbing
131 infrastructure replacement[12], we undertook a prospective study systematically
132 sampling all wastewater sites on six wards for 6-12 months alongside patient rectal
133 screening/clinical sampling, and used an anonymised electronic database to
134 characterise patient admission, ward movement, sampling profiles and culture results.
135 Multiple colonies were isolated from KPC-E-positive samples to define within-niche
136 diversity, and whole genome sequencing (Illumina) of individual isolates was used as
137 to characterise this diversity and consider modes of evolution and transmission.
138

139 **METHODS**

140 *Study setting and epidemiological data*

141 MFT is a large, regional referral centre in northwest England, UK, managing >10,000
142 patients/year. In response to the regional emergence of KPC-E[13], the Trust
143 implemented an extensive Infection Prevention and Control (IPC) programme,
144 consistent with UK CPE guidelines (published in 2014)[14]. Despite this, in April 2015
145 there was a large bla_{KPC}-*E. coli* outbreak in the cardiac unit[12], leading to its closure
146 (September 2015-January 2016) and complete refurbishment, including plumbing
147 replacement. Subsequently systematic, prospective wastewater site sampling was
148 undertaken (described further below).
149

150 To evaluate patient-level microbiological data and hospital stays, MFT electronic
151 bacteriology records were linked, based on NHS numbers, to patient administration
152 data, and anonymised. For this study, we analysed anonymised patient data from 1st
153 January 2016-31st December 2016 inclusive, and included only those patients with
154 exposure to any one of the three units (i.e. six study wards) during that time.

155

156 *Environmental sampling and laboratory processing*

157 All sink/drain/wastewater sites on six wards in three units were sampled: the cardiac
158 unit (wards 3 [W3] and 4 [W4]); the geratology unit (wards 45 [W45] and 46 [W46]);
159 and the acute medicine unit (wards AM1 and AM2) (Table S1 shows sampling site
160 designations). W3/W4 wastewater sites were sampled fortnightly on rotation from 8th
161 January 2016-28th December 2016. W45/W46 and AM1/AM2 wastewater sites were
162 similarly sampled, but from 18th July 2016-31st December 2016. Wastewater site
163 types are described in Appendix Methods.

164

165 Environmental sampling was carried out by aspirating ~20mls of wastewater from
166 sink P-traps, shower drains or toilets and performing enrichment-based culture (see
167 Appendix methods). Multiplex, real-time, quantitative PCR (for bla_{NDM}, bla_{KPC},
168 bla_{OXA-48}[12]) was performed on broths following incubation. 10µl of any bla_{KPC}-
169 positive-broth were streaked onto CPE-selective agar plates (Chromid CARBA) and
170 re-incubated (aerobically, 37°C).

171

172 *Patient sampling and laboratory processing*

173 Routinely collected patient clinical samples were processed using standard operating
174 procedures in line with UK Standards[15]. Rectal swabs were screened for CPE using
175 the Cepheid Xpert Carba-R assay or an in-house multiplex PCR (identifying bla_{KPC},
176 bla_{NDM}, and bla_{OXA-48}) from November 2014. Species identification of isolates was
177 performed by MALDI-ToF (Bruker Biotyper, Bruker MD); antimicrobial
178 susceptibility testing was performed as per EUCAST guidelines[16].

179

180 *Analyses of within-niche diversity*

181 To evaluate within-niche diversity in the environment and patients, up to five
182 different Enterobacterales colonies were individually sub-cultured; bla_{KPC}-positivity
183 in sub-cultures was re-confirmed by PCR and species identification performed by
184 MALDI-ToF prior to storage in nutrient broth at -80°C.

185

186 *DNA extraction and isolate sequencing*

187 DNA was extracted from frozen sub-culture stocks with the QuickGene kit, as per the
188 manufacturer's instructions, with an additional mechanical lysis step post-chemical
189 lysis (6m/s at 40secs x2 FastPrep, MPBio). Isolates were sequenced on the Illumina
190 HiSeq 4000 generating 150bp paired-end reads.

191

192 *Sequence read processing and analysis*

193 Sequencing reads were initially trimmed/quality controlled[17]; species identification
194 was performed with Kraken[18]. If at least 50 sequenced isolates were of a single
195 species, reads were mapped to species-specific references (Appendix Methods).
196 Variant calling and filtering was performed, and consensus fasta sequences of variant
197 calls with respect to the reference generated. Recombination-adjusted phylogenies by
198 species were created from core chromosomal single nucleotide variants
199 (SNVs)(padded to the length of the reference genome prior to analysis) using IQtree[19]
200 (flags: -m GTR+G -blmin 0.00000001 -t PARS) and then ClonalFrameML[20] (default
201 parameters). A distance of ≤400 SNVs was used as a threshold to define strains
202 (threshold used similar to[21]; Fig.S2A-E for most common species identified). For
203 detailed sub-strain analysis, sub-strains were defined on the basis of accessory

204 genome clusters characterised using Panaroo (default parameters including --clean-
205 mode strict[22]) and clustered using the R heatmap package.
206
207 Isolates were assembled using SPAdes (v3.6)[23] with default parameters. Multi-locus
208 sequence types were derived using BLASTn to relevant species-specific databases
209 hosted at https://pubmlst.org (Achtman 7-gene MLST for *E. coli*). AMR genes were
210 identified using ARIBA[24] with the CARD database[25]. bla_{KPC} genes, any Tn*4401*
211 transposon context and 5-bp target site sequences reflecting Tn*4401* insertion
212 signatures were characterised using TETyper[26], using a Tn*4401*b reference sequence,
213 a 5bp flank length and the default structural and SNV profiles.
214
215 For the analysis of *K. pneumoniae* strain 9 (see Results), putative transmission
216 networks were inferred first using SCOTTI[27] (Appendix Methods), and then
217 heuristically, linking isolates identified as clusters on the basis of accessory
218 component profiles and geographic/temporal overlap (Appendix Methods).
219
220 Sequence data have been deposited in NCBI's GenBank (BioProject accession
221 PRJNA768622).
222
223 *Statistical analyses and data visualisation*
224 Descriptive statistics were calculated in R v3.6.2 or STATA v16.1. To conduct
225 permutational analyses of variance, we used the *adonis* function from the vegan R
226 package (v2.5-7)[28] on a matrix of pairwise Gower distances based on genetic features
227 (i.e. composite of species-strain type, AMR gene profile, IS profile, and
228 Tn*4401*+flanking sequence profiles), which was calculated using the *daisy* function in

229 the R cluster package. To evaluate the impact of unit location or environmental site
230 type on the odds of CPE positivity, we used logistic regression with robust standard
231 errors clustered by environmental site sampled, using the R rms package (v6.2-0)[29].
232 Data visualisation was using the ggplot2[30] and scatterpie[31] (v.0.1.5) R packages, and
233 BioRender (www.biorender.com).
234
235 *Ethical approvals*
236 This study was performed as part of a Trust board-approved and UK Health Security
237 Agency (formerly Public Health England) long-term outbreak response, and ethics
238 approval was not required under NHS governance arrangements; all analyses were
239 undertaken using anonymised data. This was reviewed and agreed by the Joint
240 Research Office of the Oxford University Hospitals NHS Foundation Trust and the
241 University of Oxford (who oversee research governance, ethics and assurance for our
242 institution), and who deemed that neither sponsorship nor research ethics review was
243 required. Of note, sample and patient numbers represented in the
244 figures/supplementary data are study-specific labels that are not patient-identifiable.
245
246 **RESULTS**
247 *Environmental sampling revealed high prevalence of wastewater site KPC-E*
248 *colonisation*
249 349 sites across the six study wards (i.e. three units) were sampled a total of 4,488
250 times (Fig.1): Of these, 23/76 (30%) sites and 111/1,564 (7.1%) sampling events on
251 the cardiac unit were positive for KPC-E, 38/129 (29%) sites and 86/1,309 (6.6%)
252 sampling events on the acute medicine unit and 40/144 [28%] and 122/1,615 (7.6%)
253 on the geratology unit (Fig.2, top three panels). 248/349 (71.1%) sites were never

254 identified as KPC-E-positive (Fig.S3); 7/349 (2.0%) sites appeared more likely to be
255 persistently KPC-E positive (≥10 consecutive KPC-E-positive sampling events) and
256 interestingly five of these were medicines/treatment room handwash basin drains
257 likely only accessible to healthcare staff (A59, F45, F113, T43, T44; Fig.S4). The
258 proportion of positive sites did not vary significantly over time (Fig.2, bottom panel).
259 Two utility room sink drains on the cardiac unit (across W3, W4) were positive for
260 KPC-E a day after W3 reopened to patient admissions (11th January 2016) following
261 closure and complete plumbing replacement over three months; thereafter the unit
262 was rapidly re-colonised (Fig.2)[12]. Shower/bath drain sites and sluices/sluice
263 sinks/mop sinks were significantly more likely to be positive for KPC-E than others
264 (Adjusted Odds Ratio [AOR] (95% CI): 2.69 (1.44-5.01), p=0.0019 and 2.60 (1.04-
265 6.52), p=0.041, respectively), and toilet water sampling less likely to be positive
266 (AOR (95% CI): 0.28 (0.10-0.78), p=0.015); there was no evidence of a unit-
267 dependent difference (Table 1).
268
269 *Approximately one in twenty patients was affected by CPE, with most cases reflecting*
270 *rectal colonisation with KPC-producing K. pneumoniae, E. coli or E. cloacae*
271 Of 2,172 Enterobacterales positive-samples (obtained from 943 patients) and 2,000
272 with an associated carbapenem susceptibility result (172 carbapenem susceptibility
273 results missing), 399/2,000 (20.0%) were carbapenem-resistant Enterobacterales
274 (CRE; i.e. had intermediate or resistant susceptibilities) (Figs.1, S5). For 344/431
275 (79.8%) CRE isolates cultured from these 399 CRE-positive samples carbapenemase
276 PCR data were available: All contained at least one known carbapenemase resistance
277 gene, and 274 (79.7%) were KPC-E (Table S2, henceforth denoted CPE). These CPEs
278 were cultured from 183 patients (4.1% of all patients overall, 8.4% of all patients with

279 an Enterobacterales-positive culture) during the study period (median=1, IQR: 1-2,
280 range: 1-17 CPE isolates). Amongst these patients, CPE-positive samples represented
281 a median of 10% (IQR: 5-20%) of all samples taken and cultured.
282
283 The majority of confirmed CPE-positive samples were rectal screens (n=318/399
284 [79.7%]); others were non-screening clinical specimens (81/399 [20.3%] including
285 blood cultures [n=4], urine [45], respiratory [20], wound/intra-abdominal [12]) (Table
286 S3). 16/183 individuals were positive only on clinical culture (9%), 141 only on rectal
287 screen (77%), and 26 (14%) on both. By routine laboratory methods, at least 16
288 species and eight genera were represented amongst CPEs, with *K. pneumoniae*, *E.*
289 *coli* and *E. cloacae* complex predominating (364/431 [84.5%]). Notably, the
290 proportions of carbapenem-resistant cases within these top species varied (n=239/464
291 [51%] vs 88/1,083 [8%] vs 37/86 [43%] respectively; Fisher's exact test, $p<0.001$,
292 Fig.S5).
293
294 The proportion of patient CPE-positive cultures was lower than that of environmental
295 site samples, typically ~3% (Fig.3). Occasional peaks in prevalence were observed,
296 consistent with outbreaks (e.g. week 31, acute medicine unit). CPE prevalence
297 amongst patient cultures on the refurbished cardiology ward was 0% (97.5% CI: 0-
298 0.02) for the first 10 weeks post-reopening.
299
300 *Structuring of diversity by niche, with most diversity explained by the individual niche*
301 Overall, amongst successfully sequenced isolates (n=1,732) across patients and
302 environmental reservoirs, at least 13 Enterobacterales species harboured bla_{KPC}, with
303 109 unique strains represented (and a further ten species not characterised at the strain

304 level because of small numbers, see Methods). Of the 109 strains, 61 (56%) were
305 found only in the environment, 27 (24%) only in patients and 21 (19%) in both
306 patients and the environment (Fig.4); the median strain cluster size was 5 isolates
307 (range: 1-122), and the median number of unique niches (i.e. unique patients or
308 wastewater sites) affected by a strain was 2 (range: 1-23) (Fig.4A, Fig.4B). A much
309 greater proportion of all identified Tn*4401*/TSS types was observed in the
310 environment versus patients (118/121 [98%%] of types versus 9/121 [7%]
311 respectively; Fig.S6; Fisher's exact test, p<0.001), and all the isolates with >1 set of
312 TSSs were seen in environmental sites, consistent with these representing a more
313 favourable niche for Tn*4401*-associated transposition/evolution and dissemination.
314
315 For colonised niches (i.e. patient or environmental), using a metric of genomic
316 diversity based on strain type, AMR gene profile, IS profile, and Tn*4401*+ 5bp target
317 site sequence profiles, most variance was explained by the specific individual niche
318 sampled (R^2=46.0%; p<0.001), versus small but significant contributions made by
319 reservoir type (patient vs specific environmental niche; R^2=5.9%; p<0.001) and unit
320 location (R^2=3.0%; p<0.001). For the fifteen examples where exactly the same strain,
321 AMR gene profile, IS profile, plasmid profile and Tn*4401*/TSS type were seen in
322 patient and environmental niches, the temporal/geographic relationship of these
323 suggested: patient>environment transmission in 14/31 (45%) events,
324 environment>patient transmission in 2/31 (7%) events, environment>environment
325 transmission in 4/31 (13%) events, and patient>patient transmission in 11/31 (35%)
326 events (test of equality of proportions, p<0.001).
327

328 *Environmental and patient niches harbour diverse KPC-E within single samples and*
329 *over time*
330 Environmental sites showed evidence of major strain-level diversity (Fig.5): in the
331 101/349 sites from which KPC-E was successfully sequenced at any timepoint,
332 78/319 (24.4%) KPC-E-positive samples had >1 KPC-E strain, namely 52/319
333 (16.3%) with two strains, 16/319 (5.0%) with three, 9/319 (2.8%) with four and 1/319
334 (0.3%) with five strains. Over the study period, colonised environmental sites had a
335 median of 3 (range: 1-17, IQR: 2-5) KPC-E strains. 49/101 sites (49%) were
336 apparently colonised by a single KPC-E strain only, of which 38 were at single
337 sampling-points, with 11 sites showing persistent colonisation (≥2 timepoints positive
338 with the same strain).
339
340 For patient samples, based on routine laboratory microbiology, evidence of mixed-
341 species KPC-E colonisation/infection was found in 34/399 (9%) samples from 27/183
342 (15%) patients; these were predominantly *E. coli/K. pneumoniae* mixtures (n=21
343 samples; maximum two species identified). To evaluate within-host diversity with
344 genomic rather than species-level resolution, in a subset of 76 patients, strains were
345 sequenced from at least one CPE-positive sample during the study period; for 97
346 samples, multiple colonies were picked. In total 463 sequenced colonies from 97
347 samples were available for evaluation of within-patient diversity; 25/97 samples
348 (26%) had more than one strain per sample, namely 17/97 samples (18%) with two
349 KPC-E strains, 6/97 (6%) with three strains and 2/97 (2%) with four strains (Fig.5) -
350 i.e. considerably more diversity identified than based on routine laboratory species
351 identification data (26% vs 9% samples; Fisher's exact test, $p<0.001$). Longitudinal
352 switches in strain composition of KPC-E-positive samples was common, occurring in

353 11/21 (52%) KPC-E-positive longitudinal patient samples and 108/202 (53%) KPC-
354 E-positive longitudinal wastewater site samples.
355
356 *Phylogenetic analysis reveals distinct population biology for different KPC-E strains*
357 To investigate different modes of bla_{KPC} dissemination, we characterised in detail
358 several clusters, including the largest environmental-only cluster (*E. cloacae* strain 10
359 [n=122 isolates]), a large cluster involving patients and the environment (*K.*
360 *pneumoniae* strain 9 [n=116 isolates], and a cluster in which signatures of Tn*4401*
361 transposition appeared highly frequent (K. *pneumoniae* strain 11 [n=106 isolates]).
362
363 *Rapid churn of diversity in restricted environmental niches observed at the*
364 *chromosome, plasmid, IS, and AMR gene level - E. cloacae strain 10*
365 *E. cloacae* strain 10 (eclo10, ST32) appeared a restricted but persistent coloniser of
366 three co-located sink drains on a geratology ward (in a drug room clinical handwash
367 basin drain [F45] and utility sink drain [F46], and a separate treatment room clinical
368 handwash basin drain [F47]) over ~5 months (18/Jul/2016-20/Dec/2016) (Fig.6A). At
369 the core chromosomal level, strains appeared largely structured by site, with 7 SNVs
370 separating the majority of isolates in F45 from those in F47, and 14 SNVs those in
371 F46 from F45 (Fig.6A). There was evidence of rapid churn in AMR gene and IS
372 content (111 and 36 different profiles respectively; Fig.6A), and all isolates in F47
373 had acquired an additional set of IncHI2/HI2A replicons on a background of FIB/FII,
374 except for 1954277_1_4, consistent with plasmid gain/loss (Fig.6B). Although the
375 Tn*4401* type and flanking sequences were largely identical (Tn*4401*a-ATTGA-
376 ATTGA in 105/122 (86%) isolates across sinks), there was evidence of unique

377 additional Tn*4401* deletions occurring in multiple isolates at four distinct timepoints
378 in the three sites (Fig.6C).
379
380 *Rapid patient-patient transmission - K. pneumoniae strain 9*
381 *K. pneumoniae* strain 9 (kpne9, ST252) was highly related (Fig.S7; maximum 6 SNVs
382 across strain, 68 isolates within 0 SNVs of each other), and spread amongst 21
383 patients across three wards on two units (acute medicine AM1 and AM2 and
384 geratology W46), with a dense outbreak in Aug 2016, most consistent, at least
385 initially, with direct/indirect patient-patient transmission (Fig.7). bla_{KPC} was
386 consistently nested in Tn*4401*-ATTGA-ATTGA, except for one isolate which
387 acquired a Tn*4401*-associated mutation (isolate: 2216698_14, C4620T), and another
388 in which the right target site sequence (TSS) could not be identified (2035791_11).
389 Amongst the 106 isolates there was evidence of four transient plasmid replicon
390 acquisitions on a background of the stable presence of a set of IncFIB/IncFII replicons
391 (Fig.S7). Transmission probabilities inferred by SCOTTI were weak (see Appendix;
392 Fig.S8), as previously[8], but using accessory cluster profiling to attempt further
393 discrimination amongst strains suggested that there were several major transmission
394 clusters (Fig.S9), with some individuals involved in several of these (Fig.7). A single
395 shower drain site was involved (A13, serving a particular bay with four beds), which
396 was likely a seeded bystander from affected patients, but may have contributed to
397 transmission of sub-strain 3g (Fig.7).
398
399 *Cross-unit dissemination with evidence of significant Tn4401-associated* bla_{KPC}
400 *transposition - K. pneumoniae strain 11*

401 Given that Tn*4401* transposition is associated with bla_{KPC} dissemination[21], we also
402 evaluated whether the number of different Tn*4401*-TSS types was different across
403 strains, noting that the observed frequency of different Tn*4401*-TSS combinations
404 was higher in *K. pneumoniae* strain 11 (kpne11, ST11) than in other strains (Fig.S10).
405 Kpne11 represented two sub-clusters, separated by ~40SNVs (Fig.7, designated
406 "cluster 1" and "cluster 2"), both of which were isolated from patients and wastewater
407 sites across ward settings, including the cardiac unit after plumbing replacement.
408 Within both clusters of isolates there were ten different Tn*4401*a flanking signatures
409 observed, only one of which was seen in the other >1700 isolates sequenced in the
410 study (Tn*4401*a-1_ATTGA_ATTGA, the most common flanking signature), and
411 therefore most strongly consistent with the occurrence of multiple transposition
412 events within this *K. pneumoniae* strain. Similarly plasmid replicon gain/loss events
413 were observed (Col, IncN, IncR and IncFIA); all isolates contained IncFIB and IncFII
414 replicons.
415
416 **DISCUSSION**
417 This work highlights the substantial CPE diversity and flux that can occur in both
418 human and wastewater niches in healthcare settings. Many CPE strains have unique
419 features likely to make generalising results from analyses of transmission, evolution
420 and gene sharing of bla_{KPC} amongst these species very challenging. Although some
421 sites are persistently colonised with individual strains over a period of months, many
422 appear to be only transiently colonised. Sites particularly associated with CPE
423 colonisation included sluices, sluice sinks, mop sinks and shower/bath drains,
424 supported by other studies[5], and consistent with these representing a major interface
425 between human faeces and premise plumbing. Toilets were less likely to be CPE-

426 positive, perhaps because of regular flushing or particular cleaning protocols;
427 consistent with this, several studies have demonstrated that daily bleach application
428 decontaminated wastewater sites[32, 33].
429
430 At the strain-level, likely transmissions occurred from patient-patient, environment-
431 patient, patient-environment and environment-environment; however, our study
432 highlights the difficulty in accurately tracking strains across niches, and the detailed
433 and dense sampling effort required to understand within-sample diversity and
434 longitudinal flux. This difficulty has also been demonstrated in other recent work
435 investigating $bla_{\text{OXA-48}}$, where only patient-based sampling was undertaken and most
436 $bla_{\text{OXA-48}}$-associated transmission events either remained non-delineated or were only
437 weakly supported by joint phylogenetic/epidemiological inference[8]. Our study clearly
438 shows that different species and lineages represent different epidemiological risks and
439 have their own behaviours, and a focus on single, clonal strains as mediators of AMR
440 gene transmission is not relevant in this and potentially many other settings where
441 important AMR genes have become endemic.
442
443 The importance of evaluating asymptomatic CPE carriage, characterising within-host
444 diversity of strains, considering rapid acquisition/loss of strains over short
445 timeframes, and including environmental sites in transmission networks has been
446 partly demonstrated in previous studies investigating ESBL-*E. coli*[10, 34] and
447 *Enterococcus* spp.[35], but generally these features are not jointly investigated because
448 doing so is highly resource-intensive. However, without this effort, such as in studies
449 which consider only single or infrequent sampling timepoints, and do not investigate

450 asymptomatic colonisation or consider within-site diversity[36], our study highlights
451 that relevant events may easily be missed.
452
453 There are several limitations to this work. We undertook IPC-led real-time collection
454 patient isolates; sample and data capture for this process was not 24/7, and we
455 therefore only evaluated a subset of all the patient-associated CPE isolates. We were
456 not resourced to perform long-read sequencing to enable plasmid sequence
457 reconstruction, and plasmid replicon or bla_{KPC}-contig profiling is relatively low-
458 resolution[21]; we therefore avoided a detailed plasmid analysis which would be highly
459 relevant to evaluating horizontal bla_{KPC} transmission. However, even with relatively
460 low-resolution approaches, the diversity in bla_{KPC} contexts was substantial,
461 supporting high rates of HGT. A further limitation is the absence of metadata
462 capturing relevant selection pressures in each niche, including cleaning protocols for
463 environmental reservoirs and drug prescription/antimicrobial stewardship data for
464 patients. Although we only investigated one healthcare setting, the polyspecies,
465 polyclonal phenomenon we observed has been seen in other centres and with other
466 carbapenemase genes[4,8]. We focused only on Enterobacterales and may have missed
467 wider HGT of carbapenemase genes to other bacterial orders. In line with this, at least
468 23 *Pseudomonas* spp., five *Stenotrophomonas maltophilia* and two *Acinetobacter* spp.
469 that were PCR-positive for bla_{KPC} or bla_{NDM} were inadvertently picked up on rectal
470 screening.
471
472 Further optimisation of methods to evaluate multi-level genetic transmission and
473 within-sample diversity without individual colony-level characterisation is needed.
474 Holistic approaches using fully reconstructed chromosome/plasmid assemblies to

475 quantify the diverse types of genomic transmission (including clonal transmission,
476 horizontal transfer at the gene, transposon and plasmid-level) are also required to
477 better understand what facilitates emergence, selection and persistence of different
478 lineages. This information will improve our understanding of how to intervene to
479 limit the transmission of drug-resistant Enterobacterales in healthcare settings, and
480 avoid incorrect inference.

481 **Declaration of interests**

482 DWE has received lecture fees from Gilead, outside the submitted work.

483

484 **Acknowledgements**

485 We are grateful to the microbiology staff and infection prevention and control teams
486 at Manchester University NHS Foundation Trust, and to Alvaro San Millán
487 Ricardo León-Sampedro for helpful discussion.

488

489 This study was funded by the National Institute for Health Research (NIHR) Health
490 Protection Research Unit in Healthcare Associated Infections and Antimicrobial
491 Resistance (NIHR200915), a partnership between the UK Health Security Agency
492 (UKHSA) and the University of Oxford, and was supported by the NIHR Oxford
493 Biomedical Research Centre (BRC). The computational aspects of this research were
494 funded from the NIHR Oxford BRC with additional support from the Wellcome Trust
495 Core Award Grant Number 203141/Z/16/Z. The views expressed are those of the
496 author(s) and not necessarily those of the NHS, NIHR, UKHSA or the Department of
497 Health and Social Care. For the purpose of open access, the author has applied a
498 Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript
499 version arising.

500

501 DWE is a Robertson Foundation Fellow and an NIHR Oxford Biomedical Research
502 Centre Senior Fellow. NS is an Oxford Martin Fellow and an NIHR Oxford
503 Biomedical Research Centre Senior Fellow. DWC, TEAP and ASW are NIHR Senior
504 Investigators.

505

506 The Transmission of Carbapenemase-producing Enterobacteriaceae (TRACE)
507 Investigators' Group is made up of the following (listed alphabetically, includes
508 several of the authors also listed by name in the main author list): Zoie Aiken,
509 Oluwafemi Akinremi, Aiysha Ali, Julie Cawthorne, Paul Cleary, Derrick W. Crook,
510 Valerie Decraene, Andrew Dodgson, Michel Doumith, Matthew J. Ellington, Ryan
511 George, John Grimshaw, Malcolm Guiver, Robert Hill, Katie L. Hopkins, Rachel
512 Jones, Cheryl Lenney, Amy J. Mathers, Ashley McEwan, Ginny Moore, Andrew
513 Mumford, Mark Neilson, Sarah Neilson, Tim E.A. Peto, Hang T.T. Phan, Mark
514 Regan, Anna C. Seale, Nicole Stoesser, Jay Turner-Gardner, Vicky Watts, A. Sarah
515 Walker, Jimmy Walker, David H. Wyllie, William Welfare and Neil Woodford.
516
517 **REFERENCES**

518 1. Munoz-Price LS, Poirel L, Bonomo RA et al. Clinical epidemiology of the
519 global expansion of Klebsiella pneumoniae carbapenemases. *Lancet Infect Dis* 2013;
520 **13**: 785-96.
521 2. Logan LK, Weinstein RA. The Epidemiology of Carbapenem-Resistant
522 Enterobacteriaceae: The Impact and Evolution of a Global Menace. *J Infect Dis* 2017;
523 **215**: S28-S36.
524 3. Nordmann P, Dortet L, Poirel L. Carbapenem resistance in
525 Enterobacteriaceae: here is the storm! *Trends Mol Med* 2012; **18**: 263-72.
526 4. Mathers AJ, Cox HL, Kitchel B et al. Molecular dissection of an outbreak of
527 carbapenem-resistant enterobacteriaceae reveals Intergenus KPC carbapenemase
528 transmission through a promiscuous plasmid. *MBio* 2011; **2**: e00204-11.

529 5. Kizny Gordon AE, Mathers AJ, Cheong EY et al. Is the hospital water
530 environment a reservoir for carbapenem-resistant organisms causing hospital-acquired
531 infections? A systematic review of the literature. *Clin Infect Dis* 2017.
532 6. Bleichenbacher S, Stevens MJA, Zurfluh K et al. Environmental
533 dissemination of carbapenemase-producing Enterobacteriaceae in rivers in
534 Switzerland. *Environ Pollut* 2020; **265**.
535 7. Cahill N, O'Connor L, Mahon B et al. Hospital effluent: A reservoir for
536 carbapenemase-producing Enterobacterales? *The Science of the total environment*
537 2019; **672**: 618-24.
538 8. León-Sampedro R, DelaFuente J, Díaz-Agero C et al. Pervasive transmission
539 of a carbapenem resistance plasmid in the gut microbiota of hospitalized patients. *Nat*
540 *Microbiol* 2021; **6**: 606-16.
541 9. Weingarten RA, Johnson RC, Conlan S et al. Genomic Analysis of Hospital
542 Plumbing Reveals Diverse Reservoir of Bacterial Plasmids Conferring Carbapenem
543 Resistance. *MBio* 2018; **9**.
544 10. Stoesser N, Sheppard AE, Moore CE et al. Extensive Within-Host Diversity in
545 Fecally Carried Extended-Spectrum-Beta-Lactamase-Producing Escherichia coli
546 Isolates: Implications for Transmission Analyses. *J Clin Microbiol* 2015; **53**: 2122-
547 31.
548 11. Ludden C, Coll F, Gouliouris T et al. Defining nosocomial transmission of
549 Escherichia coli and antimicrobial resistance genes: a genomic surveillance study. *The*
550 *Lancet Microbe* 2021; **2**: e472-e80.
551 12. Decraene V, Phan HTT, George R et al. A Large, Refractory Nosocomial
552 Outbreak of Klebsiella pneumoniae Carbapenemase-Producing Escherichia coli

553 Demonstrates Carbapenemase Gene Outbreaks Involving Sink Sites Require Novel
554 Approaches to Infection Control. *Antimicrob Agents Chemother* 2018; **62**.
555 13. Public Health England. Carbapenemase-producing Enterobacteriaceae:
556 laboratory confirmed cases, 2003 to 2015. London: Public Health England, 2016.
557 14. Public Health England. Carbapenemase-producing Enterobacteriaceae: early
558 detection, management and control toolkit for acute trusts. 2014.
559 15. UK Standards for Microbiology Investigations.
560 https://www.gov.uk/government/collections/standards-for-microbiology-
561 investigations-smi. Last updated: 1st July 2021.
562 16. The European Committee on Antimicrobial Susceptibility Testing. Breakpoint
563 tables for interpretation of MICs and zone diameters. 2017.
564 17. Stoesser N, Giess A, Batty EM et al. Genome sequencing of an extended
565 series of NDM-producing Klebsiella pneumoniae isolates from neonatal infections in
566 a Nepali hospital characterizes the extent of community- versus hospital-associated
567 transmission in an endemic setting. *Antimicrob Agents Chemother* 2014; **58**: 7347-57.
568 18. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence
569 classification using exact alignments. *Genome Biol* 2014; **15**: R46.
570 19. Nguyen LT, Schmidt HA, von Haeseler A et al. IQ-TREE: a fast and effective
571 stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol*
572 2015; **32**: 268-74.
573 20. Didelot X, Wilson DJ. ClonalFrameML: efficient inference of recombination
574 in whole bacterial genomes. *PLoS Comput Biol* 2015; **11**: e1004041.
575 21. Sheppard AE, Stoesser N, Wilson DJ et al. Nested Russian Doll-Like Genetic
576 Mobility Drives Rapid Dissemination of the Carbapenem Resistance Gene blaKPC.
577 *Antimicrob Agents Chemother* 2016; **60**: 3767-78.

578 22. Tonkin-Hill G, MacAlasdair N, Ruis C et al. Producing polished prokaryotic
579 pangenomes with the Panaroo pipeline. *Genome Biol* 2020; **21**: 180.
580 23. Bankevich A, Nurk S, Antipov D et al. SPAdes: a new genome assembly
581 algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012; **19**: 455-
582 77.
583 24. Hunt M, Mather AE, Sánchez-Busó L et al. ARIBA: rapid antimicrobial
584 resistance genotyping directly from sequencing reads. *Microb Genom* 2017; **3**:
585 e000131.
586 25. Jia B, Raphenya AR, Alcock B et al. CARD 2017: expansion and model-
587 centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids*
588 *Res* 2017; **45**: D566-D73.
589 26. Sheppard AE, Stoesser N, German-Mesner I et al. TETyper: a bioinformatic
590 pipeline for classifying variation and genetic contexts of transposable elements from
591 short-read whole-genome sequencing data. *Microb Genom* 2018; **4**.
592 27. De Maio N, Wu CH, Wilson DJ. SCOTTI: Efficient Reconstruction of
593 Transmission within Outbreaks with the Structured Coalescent. *PLoS Comput Biol*
594 2016; **12**: e1005130.
595 28. Oksanen JB, Blanchet, FG, Friendly M, Kindt R, Legendre P, McGlinn D,
596 Minchin PR, O'Hara RB, Simpson GL, Solymos P, Henry M, Stevens H, Szoecs E,
597 Wagner H. vegan: Community Ecology Package. 2020.
598 29. Harrell Jr FE. rms: Regression Modeling Strategies. 2021.
599 30. Wickham H. ggplot2: Elegant Graphics for Data Analysis. 2016.
600 31. Yu G. scatterpie: Scatter Pie Plot. 2020.

601 32. Valentin AS, Santos SD, Goube F et al. A prospective multicentre surveillance
602 study to investigate the risk associated with contaminated sinks in the intensive care
603 unit. *Clin Microbiol Infect* 2021; **27**: 1347.e9-.e14.
604 33. Heireman L, Hamerlinck H, Vandendriessche S et al. Toilet drain water as a
605 potential source of hospital room-to-room transmission of carbapenemase-producing
606 Klebsiella pneumoniae. *The Journal of hospital infection* 2020; **106**: 232-9.
607 34. Kantele A, Kuenzli E, Dunn SJ et al. Dynamics of intestinal multidrug-
608 resistant bacteria colonisation contracted by visitors to a high-endemic setting: a
609 prospective, daily, real-time sampling study. *The Lancet Microbe* 2021; **2**: e151-e8.
610 35. Gouliouris T, Coll F, Ludden C et al. Quantifying acquisition and transmission
611 of Enterococcus faecium using genomic surveillance. *Nat Microbiol* 2021; **6**: 103-11.
612 36. Qiao F, Wei L, Feng Y et al. Handwashing Sink Contamination and
613 Carbapenem-resistant Klebsiella Infection in the Intensive Care Unit: A Prospective
614 Multicenter Study. *Clin Infect Dis* 2020; **71**: S379-s85.
615

616 **TABLES**

617 **Table 1. Association of site type and unit location with *bla*$_{KPC}$-Enterobacterales**
618 **(KPC-E) positivity in environmental samples.** Outputs of a multivariable logistic
619 regression model with robust standard errors clustered by environmental site type and
620 unit location; no other variables were considered in the model. P-values <0.05 were
621 considered significant and are highlighted in bold. No interactions with an interaction
622 Wald p <0.05 were observed.

623

Environmental site	**Number of sites**	**Number of sampling events KPC-E-positive (% of total sampling events)**	**Number of sampling events KPC-E-negative**	**OR (95% CI)**	**p**
Environmental site type					
Clinical handwash basin drain	96	79 (5.9%)	1270	Reference	
Handwash basin drain	85	52 (4.6%)	1077	0.99 (0.51-1.93)	0.9735
Shower/bath drain	59	100 (19.2%)	421	2.69 (1.44-5.01)	**0.0019**
Sluice/sluice	16	47 (17.9%)	215	2.60 (1.04-	**0.0410**

sink/mop sink				6.52)	
Toilet	74	5 (0.5%)	931	0.28 (0.10-0.78)	**0.0152**
Utility sink	19	36 (12.4%)	255	1.32 (0.47-3.68)	0.5965
Environmental site location					
Acute medicine unit	129	86 (6.6%)	1223	Reference	
Cardiac unit	76	111 (7.0%)	1453	1.07 (0.57-2.00)	0.8412
Geratology unit	144	122 (7.6%)	1493	0.88 (0.52-1.48)	0.6243

624 **FIGURE LEGENDS**

625 **Figure 1. Sampling and sequencing flowchart for the study.**

626

627 **Figure 2. *bla*$_{KPC}$-Enterobacterales (KPC-E)-positive environmental sites by week**
628 **in 2016 and hospital unit.** The number of environmental site samples performed and
629 KPC-E-positive samples by week of sampling, stratified by unit (top three panels),
630 and the proportion of KPC-E-positive sites by week of sampling, again stratified by
631 unit (bottom panel).

632

633 **Figure 3. Carbapenem-susceptible Enterobacterales (CSE) and carbapenem-**
634 **resistant-Enterobacterales (CRE) patient cultures by week in 2016 and hospital**
635 **unit.** Top three panels show counts of Enterobacterales culture-positive samples
636 stratified by unit and carbapenem susceptibility; bottom panel shows the proportion of
637 all cultures taken that were CRE, again stratified by unit.

638

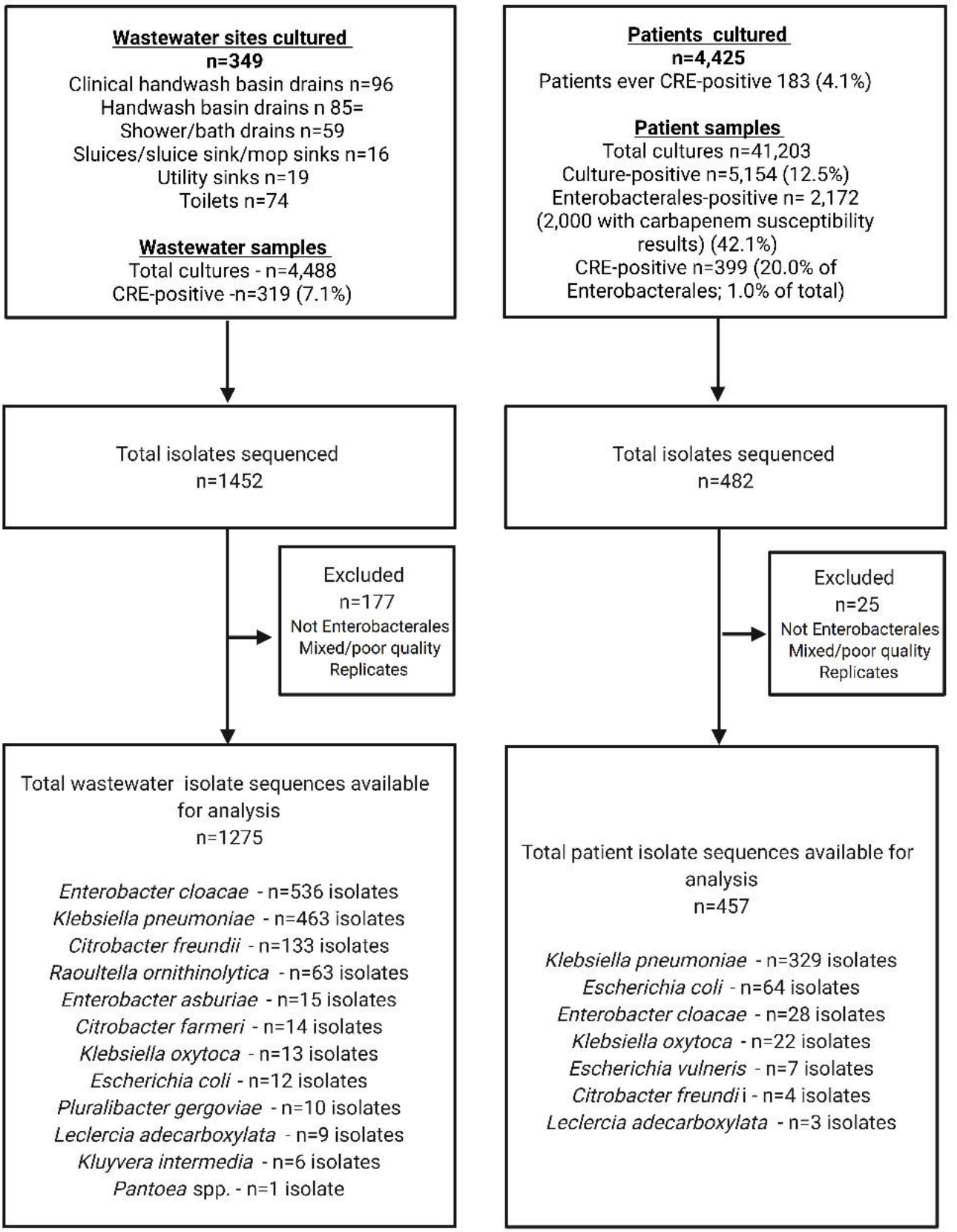
639 **Figure 4. Number of unique patient/environmental niches colonised by common**
640 ***bla*$_{KPC}$ Enterobacterales species/strains.** A) Includes species/strains identified in
641 both patients and wastewater niches during the study (if >50 isolates of a species were
642 identified [see Methods]). Orange circles denote wastewater niches, black circles
643 unique patients. B) Includes species/strains identified only in patients (left panel) and
644 only in wastewater sites (right panel). For *Raoultella ornithinolytica*, six strains were
645 identified and these were observed in wastewater sites only (data not shown).

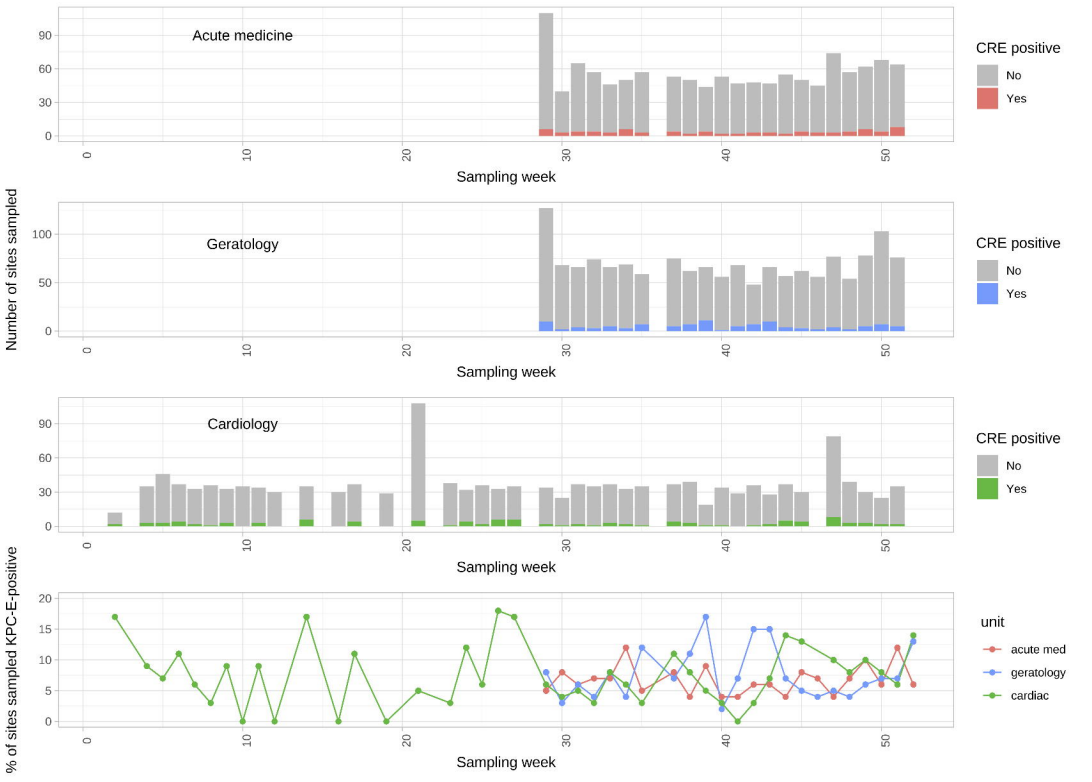
646

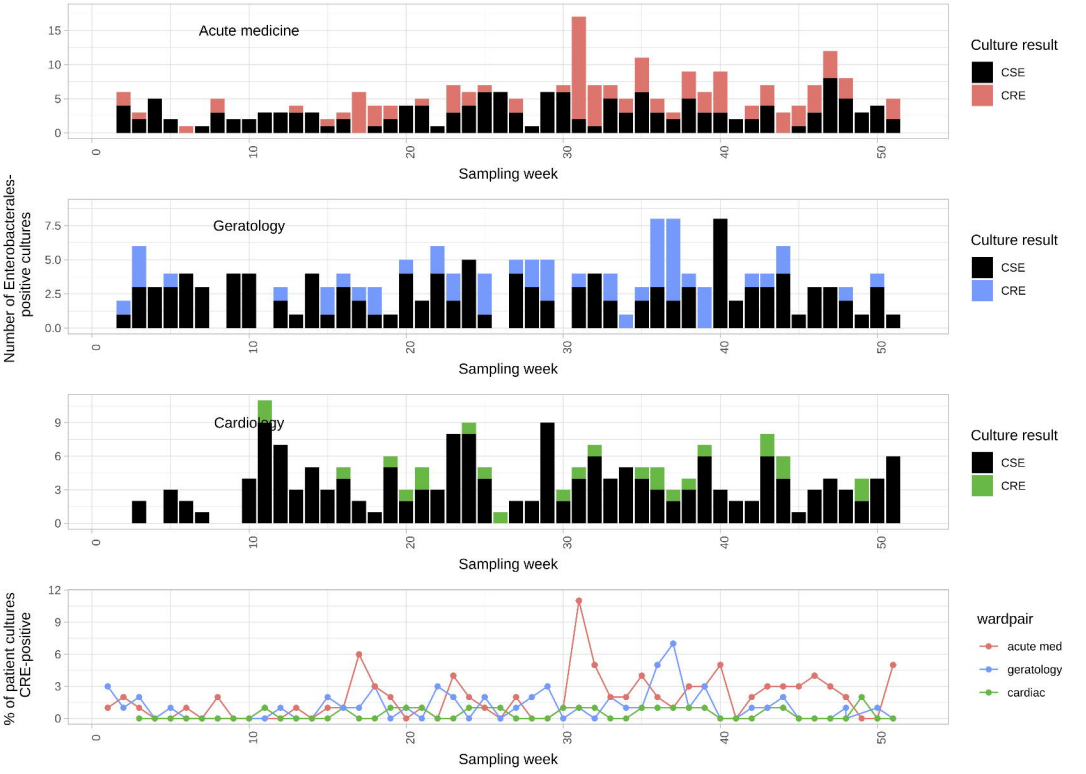
647 **Figure 5.** Top panel: Within-patient KPC-E diversity. Pies reflect the proportion of
648 strains within a sample that were identified using whole genome sequencing, with

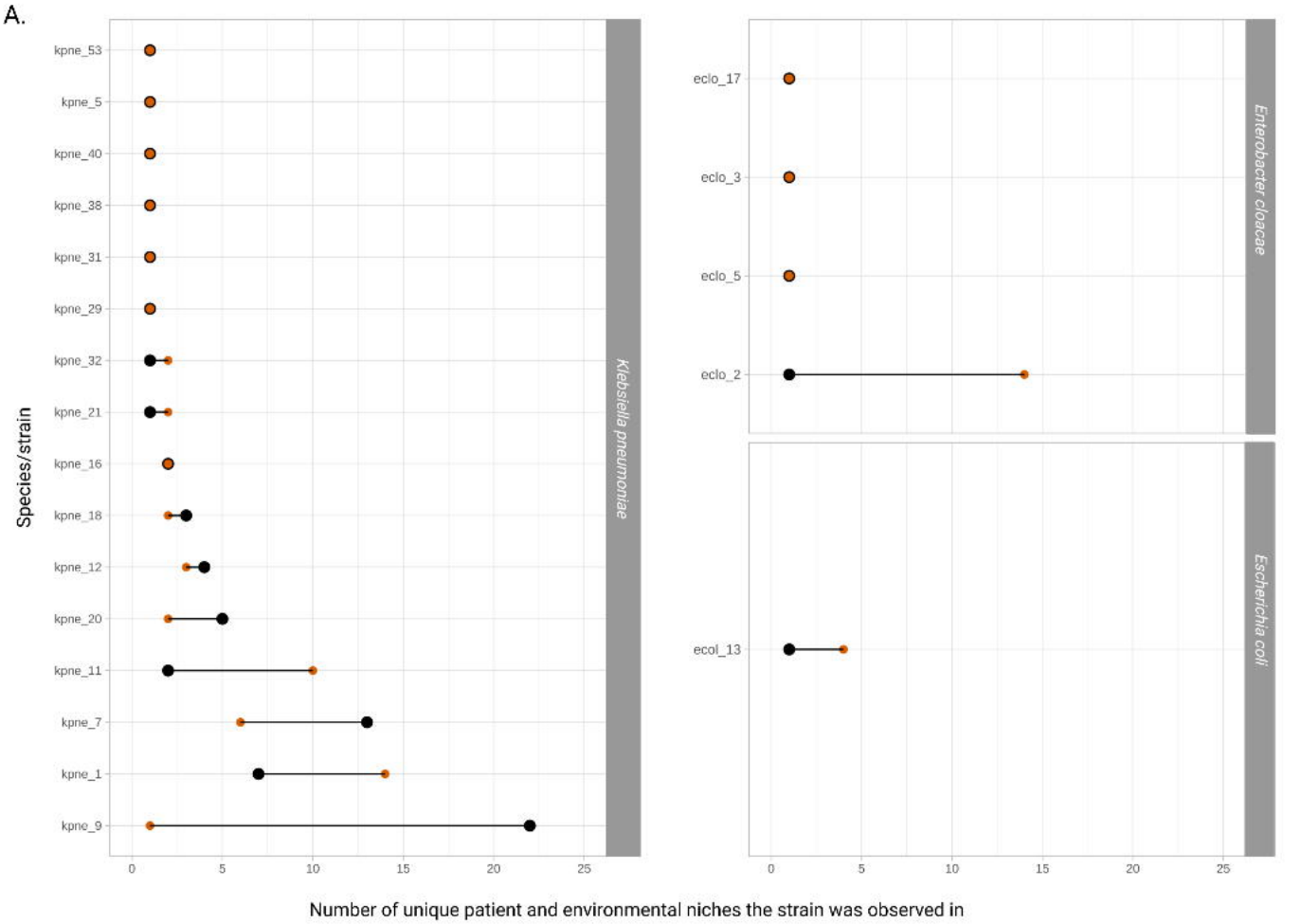
649 evaluation of diversity at the species-strain level. “cfre” = *Citrobacter freundii*, “eclo”
650 = *Enterobacter cloacae*, “ecol” = *Escherichia coli*, “evul” = *Escherichia vulnificus*,
651 “koxy” = *Klebsiella oxytoca*, “kpne”= *Klebsiella pneumoniae*, “lecl” = *Leclercia*
652 *adecarboxylata*. Numbers after species abbreviations denote unique strains. Bottom
653 panel: Within-wastewater sample KPC-E diversity. Pies reflect the proportion of
654 strains within a sample that were identified using whole genome sequencing, with
655 evaluation of diversity at the species-strain level. “aher” = *Escherichia hermanii*,
656 “cfar” = *Citrobacter farmeri*, “easb” = *Enterobacter asburiae*, “pant” = *Pantoea* spp.,
657 “plur” = *Pluralibacter gergoviae*; other assignations as for top panel.
658
659 **Figure 6**. A) Recombination-adjusted phylogeny for *E. cloacae* strain 10. Main
660 annotations as described in the legends as part of the figure. In the panel to the right
661 of the phylogeny, AMR profile, IS profile, plasmid replicon profile and Tn*4401*/target
662 site sequences (TSSs) are represented, with different colours distinguishing different
663 versions of each. Three isolates in F47 and one isolate in F46 clustered with the F45-
664 associated isolates; sampling dates suggest that these represented transmission events
665 from F45. The scale bar is in SNVs. B) Represents plasmid replicon profiles in the
666 isolates obtained from the environmental site samples over time, and C) similarly
667 represents the Tn*4401*/TSS profiles with additional deletions observed annotated as
668 an overlay for isolates typed as containing the “unknown-1_ATTGA_ATTGA
669 variants” using TETyper.
670
671 **Figure 7.** Putative transmission network for sub-lineages of *K. pneumoniae* strain 9,
672 defined by accessory genome clustering (clusters 1, 2, 3a-h). Horizontal bars
673 represent admission episodes to study wards for each patient (pale blue=W45,

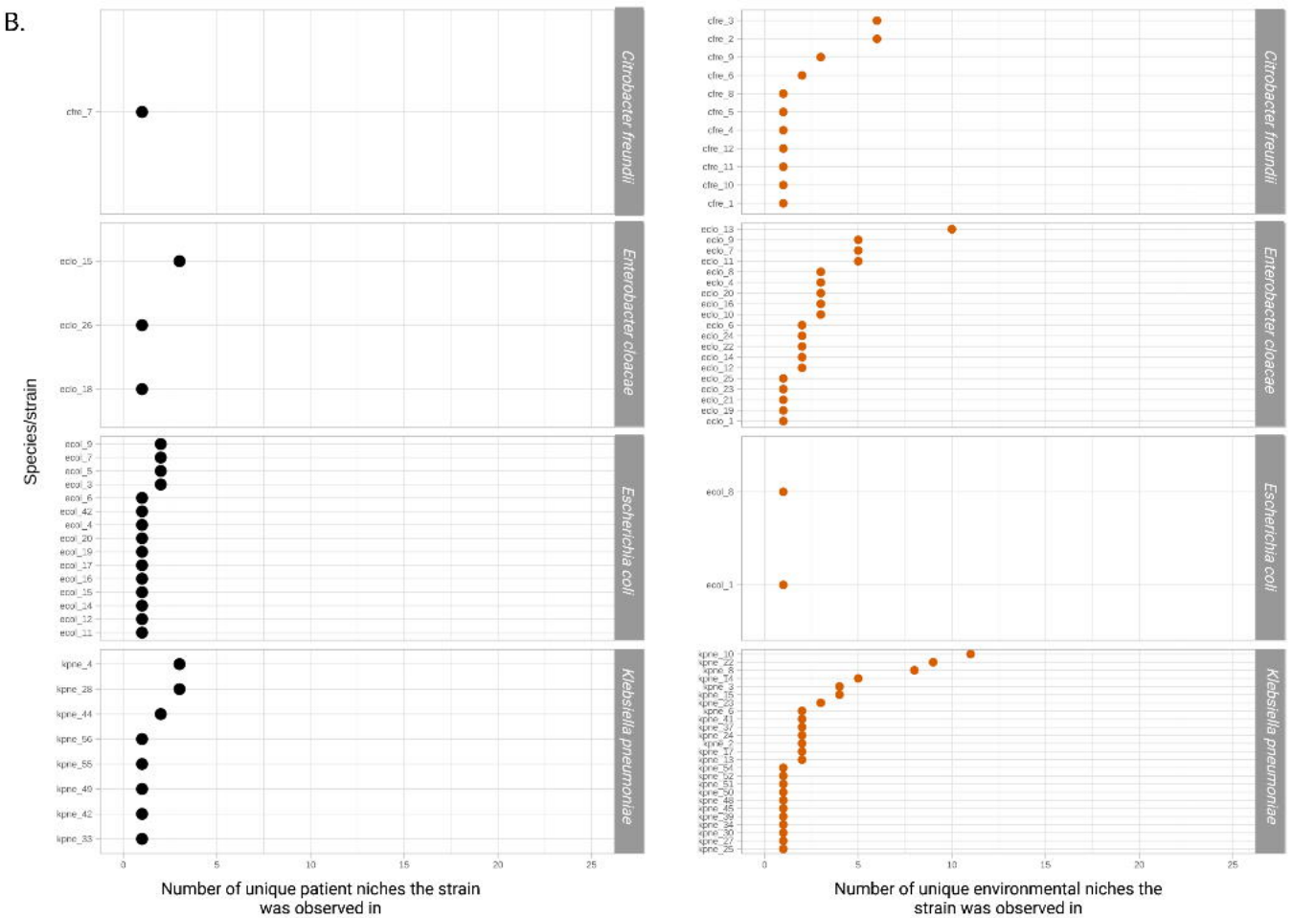
674 geratology; pale pink=AM1, acute medicine; pale purple=AM2, acute medicine; A13
675 represents the only shower site involved. Empty black triangles denote CPE-negative
676 rectal screens on patients. Arrows denote possible transmission links (dashed arrows
677 denote events that have no geographic overlap or could be associated with two links
678 based on negative screen, geographic location and timing). For clarity, arrows are not
679 drawn for clusters 3g and 3h, which affected the most patients - these are shown as an
680 inset panel at the bottom right.
681
682 **Figure 8.** Left-hand panel shows a recombination-corrected phylogeny for *K.*
683 *pneumoniae* strain 11, designating the two identified clusters, 1 and 2, and annotated
684 with isolate names, unit, collection date, niche, and Tn*4401*-TSS type. Right-hand
685 panels reflect the plasmid replicon and Tn*4401*-TSS diversity within sequenced
686 isolates in given niches over time for each chromosomally genetically-related cluster,
687 highlighting features consistent with clonal transmission, Tn*4401* transposition events
688 within isolates (Tn*4401*a, change in inner dot colour with consistent outer ring
689 colour), deletions in Tn*4401*a (to "unknown-a"), plasmid replicon loss/gain events
690 (change in outer ring colour).

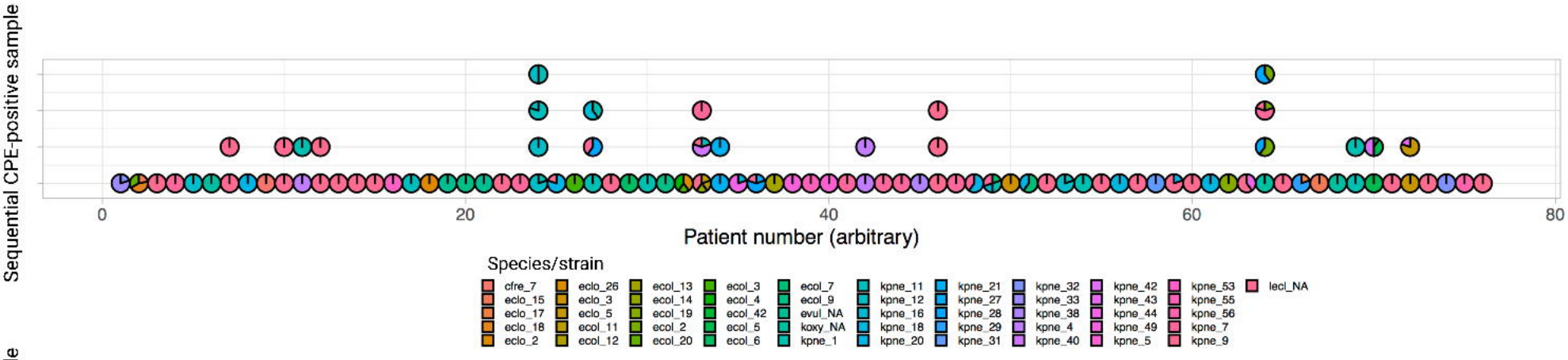

Wastewater sites cultured
n=349
Clinical handwash basin drains n=96
Handwash basin drains n 85=
Shower/bath drains n=59
Sluices/sluice sink/mop sinks n=16
Utility sinks n=19
Toilets n=74
Wastewater samples
Total cultures - n=4,488
CRE-positive -n=319 (7.1%)
Patients cultured
n=4,425
Patients ever CRE-positive 183 (4.1%)
Patient samples
Total cultures n=41,203
Culture-positive n=5,154 (12.5%)
Enterobacterales-positive n= 2,172
(2,000 with carbapenem susceptibility results) (42.1%)
CRE-positive n=399 (20.0% of Enterobacterales; 1.0% of total)
Total isolates sequenced
n=1452
Total isolates sequenced
n=482
Excluded
n=177
Not Enterobacterales
Mixed/poor quality
Replicates
Excluded
n=25
Not Enterobacterales
Mixed/poor quality
Replicates
Total wastewater isolate sequences available for analysis
n=1275
Enterobacter cloacae - n=536 isolates
Klebsiella pneumoniae - n=463 isolates
Citrobacter freundii - n=133 isolates
Raoultella ornithinolytica - n=63 isolates
Enterobacter asburiae - n=15 isolates
Citrobacter farmeri - n=14 isolates
Klebsiella oxytoca - n=13 isolates
Escherichia coli - n=12 isolates
Pluralibacter gergoviae - n=10 isolates
Leclercia adecarboxylata - n=9 isolates
Kluyvera intermedia - n=6 isolates
Pantoea spp. - n=1 isolate
Total patient isolate sequences available for analysis
n=457
Klebsiella pneumoniae - n=329 isolates
Escherichia coli - n=64 isolates
Enterobacter cloacae - n=28 isolates
Klebsiella oxytoca - n=22 isolates
Escherichia vulneris - n=7 isolates
Citrobacter freundii - n=4 isolates
Leclercia adecarboxylata - n=3 isolates

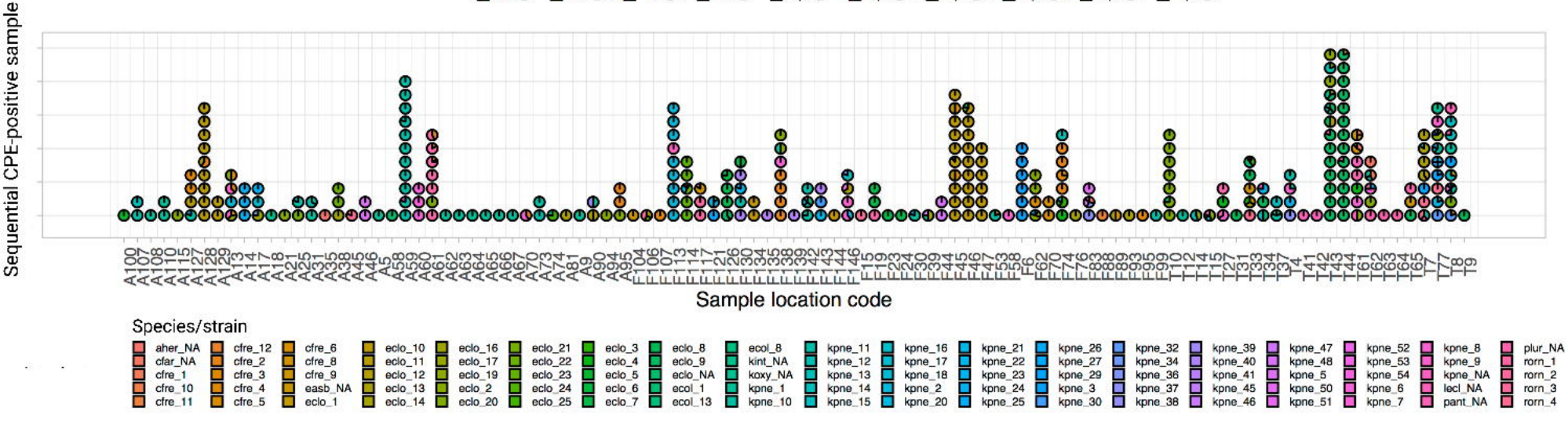

Acute medicine
Geratology
Cardiology
CRE positive
No
Yes
Number of sites sampled
Sampling week
% of sites sampled KPC-E-positive
unit
acute med
geratology
cardiac

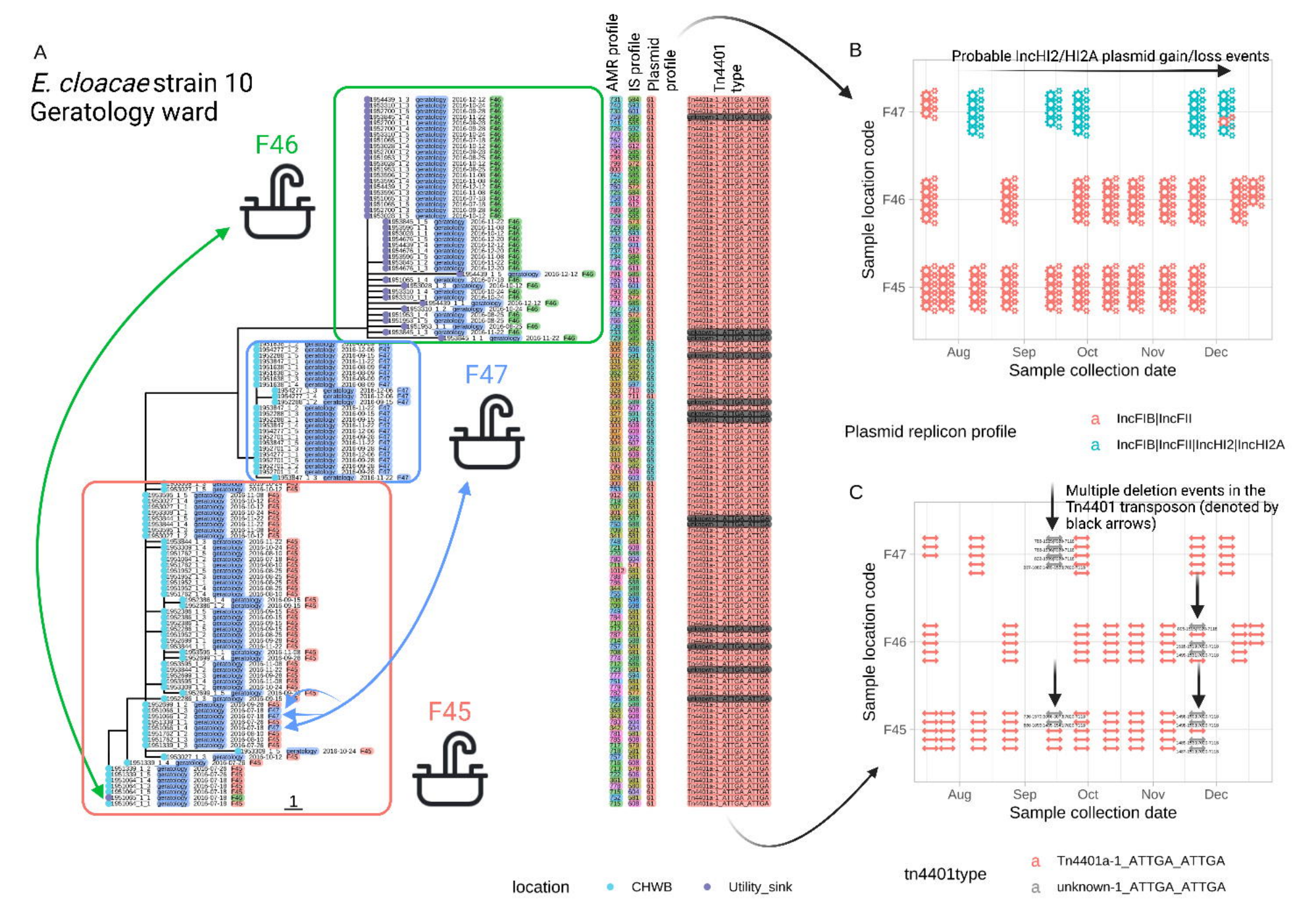

Acute medicine
Geratology
Cardiology
Number of Enterobacterales-positive cultures
% of patient cultures CRE-positive
Sampling week
Culture result
CSE
CRE
wardpair
acute med
geratology
cardiac

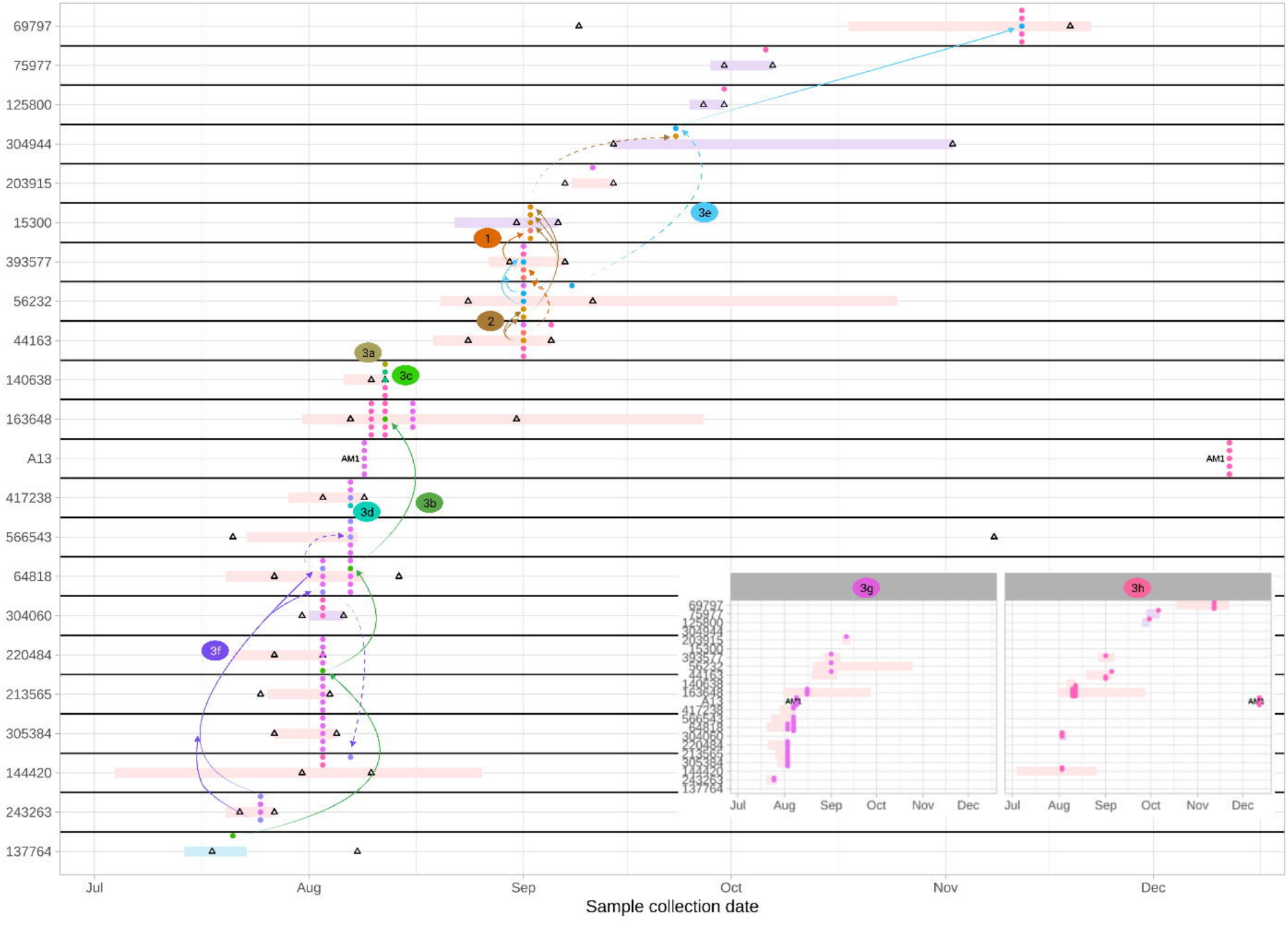

A.
Species/strain
kpne_53
kpne_5
kpne_40
kpne_38
kpne_31
kpne_29
kpne_32
kpne_21
kpne_16
kpne_18
kpne_12
kpne_20
kpne_11
kpne_7
kpne_1
kpne_9
Klebsiella pneumoniae
eclo_17
eclo_3
eclo_5
eclo_2
Enterobacter cloacae
ecol_13
Escherichia coli
0
5
10
15
20
25
Number of unique patient and environmental niches the strain was observed in

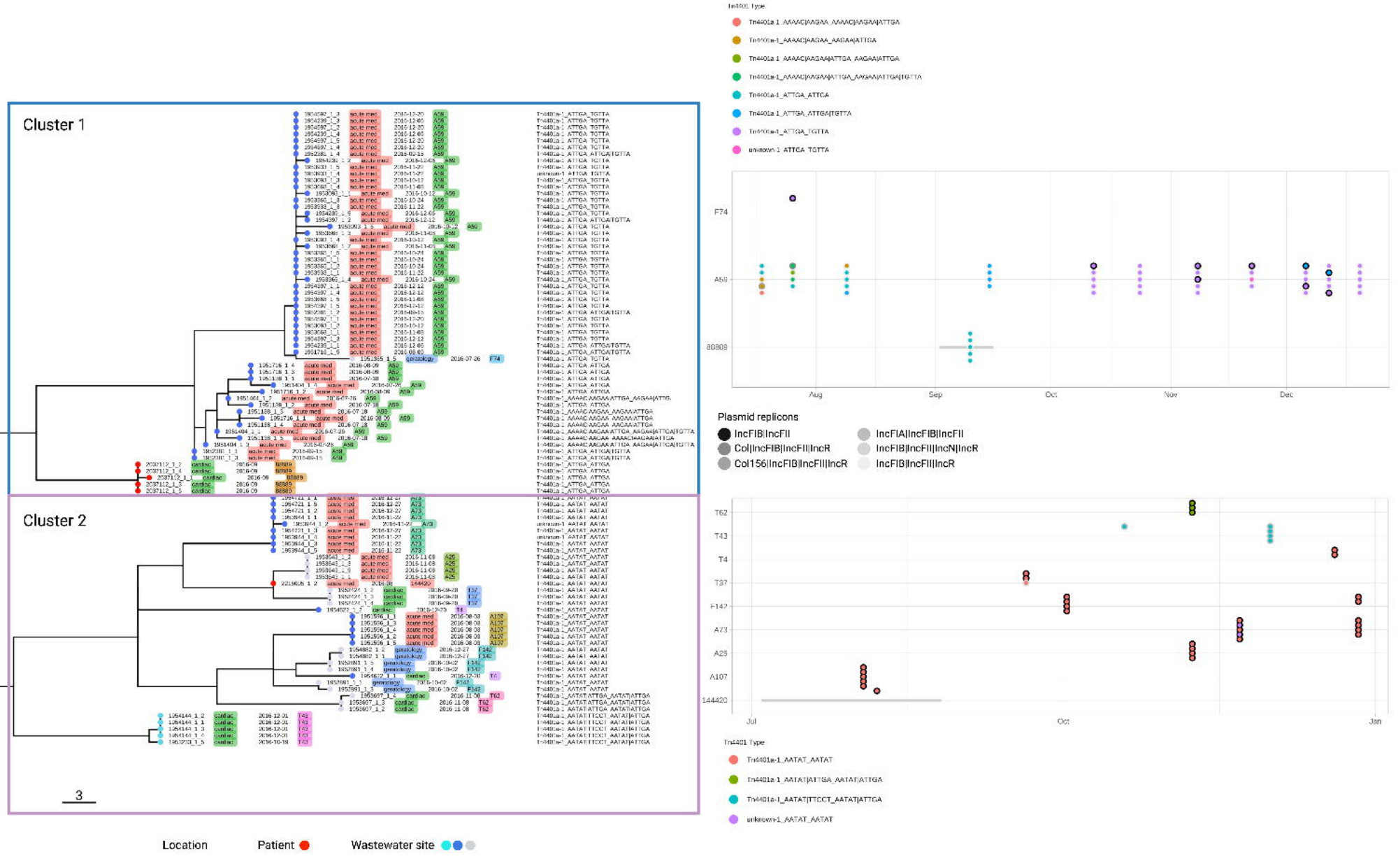

B.
Species/strain
cfre_7
Citrobacter freundii
eclo_15
eclo_26
eclo_18
Enterobacter cloacae
ecol_9
ecol_7
ecol_5
ecol_3
ecol_6
ecol_42
ecol_4
ecol_20
ecol_19
ecol_17
ecol_16
ecol_15
ecol_14
ecol_12
ecol_11
Escherichia coli
kpne_4
kpne_28
kpne_44
kpne_56
kpne_55
kpne_40
kpne_42
kpne_33
Klebsiella pneumoniae
0
5
10
15
20
25
Number of unique patient niches the strain was observed in
cfre_3
cfre_2
cfre_9
cfre_6
cfre_8
cfre_5
cfre_4
cfre_12
cfre_11
cfre_10
cfre_1
Citrobacter freundii
eclo_13
eclo_9
eclo_7
eclo_11
eclo_8
eclo_4
eclo_20
eclo_16
eclo_10
eclo_6
eclo_24
eclo_22
eclo_14
eclo_12
eclo_25
eclo_23
eclo_21
eclo_19
eclo_1
Enterobacter cloacae
ecol_8
ecol_1
Escherichia coli
kpne_10
kpne_22
kpne_8
kpne_14
kpne_3
kpne_15
kpne_23
kpne_6
kpne_41
kpne_37
kpne_24
kpne_2
kpne_17
kpne_13
kpne_54
kpne_32
kpne_51
kpne_50
kpne_46
kpne_45
kpne_39
kpne_34
kpne_30
kpne_27
kpne_25
Klebsiella pneumoniae
0
5
10
15
20
25
Number of unique environmental niches the strain was observed in


Sequential CPE-positive sample
0
20
40
60
80
Patient number (arbitrary)
Species/strain
cfre_7
eclo_15
eclo_17
eclo_18
eclo_2
eclo_26
eclo_3
eclo_5
ecol_11
ecol_12
ecol_13
ecol_14
ecol_19
ecol_2
ecol_20
ecol_3
ecol_4
ecol_42
ecol_5
ecol_6
ecol_7
ecol_9
evul_NA
koxy_NA
kpne_1
kpne_11
kpne_12
kpne_16
kpne_18
kpne_20
kpne_21
kpne_27
kpne_28
kpne_29
kpne_31
kpne_32
kpne_33
kpne_38
kpne_4
kpne_40
kpne_42
kpne_43
kpne_44
kpne_49
kpne_5
kpne_53
kpne_55
kpne_56
kpne_7
kpne_9
lecl_NA


Sequential CPE-positive sample
Sample location code
Species/strain
aher_NA
cfar_NA
cfre_1
cfre_10
cfre_11
cfre_12
cfre_2
cfre_3
cfre_4
cfre_5
cfre_6
cfre_8
cfre_9
easb_NA
eclo_1
eclo_10
eclo_11
eclo_12
eclo_13
eclo_14
eclo_16
eclo_17
eclo_19
eclo_2
eclo_20
eclo_21
eclo_22
eclo_23
eclo_24
eclo_25
eclo_3
eclo_4
eclo_5
eclo_6
eclo_7
eclo_8
eclo_9
eclo_NA
ecol_1
ecol_13
ecol_8
kint_NA
koxy_NA
kpne_1
kpne_10
kpne_11
kpne_12
kpne_13
kpne_14
kpne_15
kpne_16
kpne_17
kpne_18
kpne_2
kpne_20
kpne_21
kpne_22
kpne_23
kpne_24
kpne_25
kpne_26
kpne_27
kpne_29
kpne_3
kpne_30
kpne_32
kpne_34
kpne_36
kpne_37
kpne_38
kpne_39
kpne_40
kpne_41
kpne_45
kpne_46
kpne_47
kpne_48
kpne_5
kpne_50
kpne_51
kpne_52
kpne_53
kpne_54
kpne_6
kpne_7
kpne_8
kpne_9
kpne_NA
lecl_NA
pant_NA
plur_NA
rorn_1
rorn_2
rorn_3
rorn_4


A
E. cloacae strain 10
Geratology ward
F46
F47
F45
1
AMR profile
IS profile
Plasmid profile
Tn4401 type
B
Probable IncHI2/HI2A plasmid gain/loss events
Sample location code
F47
F46
F45
Aug
Sep
Oct
Nov
Dec
Sample collection date
Plasmid replicon profile
a IncFIB|IncFII
a IncFIB|IncFII|IncHI2|IncHI2A
C
Multiple deletion events in the Tn4401 transposon (denoted by black arrows)
Sample location code
F47
F46
F45
Aug
Sep
Oct
Nov
Dec
Sample collection date
tn4401type
a Tn4401a-1_ATTGA_ATTGA
a unknown-1_ATTGA_ATTGA
location
CHWB
Utility_sink


69797
75977
125800
304944
203915
15300
393577
56232
44163
140638
163648
A13
417238
566543
64818
304060
220484
213565
305384
144420
243263
137764
1
2
3a
3b
3c
3d
3e
3f
3g
3h
AM1
Jul
Aug
Sep
Oct
Nov
Dec
Sample collection date


Cluster 1
Cluster 2
3
Location
Patient
Wastewater site
Tn4001 Type
Tn4001a-1_AAAAC|AAGAA_AAAAC|AAGAA|ATTGA
Tn4001a-1_AAAAC|AAGAA_AAGAA|ATTGA
Tn4001a-1_AAAAC|AAGAA|ATTGA_AAGAA|ATTGA
Tn4001a-1_AAAAC|AAGAA|ATTGA_AAGAA|ATTGA|TGTTA
Tn4001a-1_ATTGA_ATTGA
Tn4001a-1_ATTGA_ATTGA|TGTTA
Tn4001a-1_ATTGA_TGTTA
unknown-1_ATTGA_TGTTA
F74
A59
80809
Aug
Sep
Oct
Nov
Dec
Plasmid replicons
IncFIB|IncFII
Col|IncFIB|IncFII|IncR
Col156|IncFIB|IncFII|IncR
IncFIA|IncFIB|IncFII
IncFIB|IncFII|IncN|IncR
IncFIB|IncFII|IncR
T62
T43
T4
T37
F142
A73
A25
A107
144420
Jul
Oct
Jan
Tn4401 Type
Tn4401a-1_AATAT_AATAT
Tn4401a-1_AATAT|ATTGA_AATAT|ATTGA
Tn4401a-1_AATAT|TTCCT_AATAT|ATTGA
unknown-1_AATAT_AATAT