

1 Metagenomic analysis of antimicrobial resistance genes in wastewaters in Benin and Burkina Faso indicates a  
2 serious health risk from untreated hospital wastewaters in low-income countries

3

4 Melina A. Markkanen<sup>a#</sup>, Kaisa Haukka<sup>a, f</sup>, Katariina M. M. Pärnänen<sup>a</sup>, Victorien Tamegnon Dougnon<sup>b</sup>, Isidore  
5 Juste O. Bonkoungou<sup>c</sup>, Zakaria Garba<sup>d</sup>, Halidou Tinto<sup>d</sup>, Anniina Sarekoski<sup>a</sup>, Antti Karkman<sup>a</sup>, Anu Kantele<sup>c, f\*</sup>,  
6 Marko P. J. Virta<sup>a#\*</sup>

7

8 <sup>a</sup> Department of Microbiology, University of Helsinki, Finland

9 <sup>b</sup> Research Unit in Applied Microbiology and Pharmacology of Natural Substances, Polytechnic School of  
10 Abomey-Calavi, University of Abomey-Calavi, Benin

11 <sup>c</sup> Department of Biochemistry and Microbiology, University Joseph KI-ZERBO, Ouagadougou, Burkina Faso

12 <sup>d</sup> Clinical Research Unit Nanoro, Institute for Research in Health Sciences, National Center for Scientific and  
13 Technological Research, Burkina Faso

14 <sup>e</sup> Meilahti Infectious Diseases and Vaccine Research Center, MeiVac, Department of Infectious Diseases,  
15 University of Helsinki and Helsinki University Hospital, Helsinki, Finland

16 <sup>f</sup> Human Microbiome Research Program, Faculty of Medicine, University of Helsinki, Helsinki, Finland

17

18 Corresponding authors:

19 # Marko P. J. Virta, [marko.virta@helsinki.fi](mailto:marko.virta@helsinki.fi)

20 # Melina A. Markkanen, [melina.markkanen@helsinki.fi](mailto:melina.markkanen@helsinki.fi)

21 \* Equal contribution

22

23 Running title

24 HWW resistomes in Benin, Burkina Faso, and Finland

25

26 Abstract

27

28 Antimicrobial resistance (AMR) is one of the greatest global threats to human health, but substantial gaps in AMR  
29 data exist in West African countries. We explored the presence of antimicrobial resistance genes (ARGs) in the  
30 hospital wastewaters (HWW) of nine hospitals in Benin and Burkina Faso and, for comparison, four hospitals in  
31 Finland using shotgun metagenomic sequencing. The highest relative sum abundance of ARGs in HWWs was  
32 observed in Benin and the lowest in Finland. HWW resistomes were more similar to each other in Benin and  
33 Burkina Faso than in Finland. Different carbapenemases were detected in varying abundances, especially in  
34 HWWs from Burkina Faso and Finland. The most widespread carbapenemase gene in the Beninese hospitals,  
35 *bla*GES, was also found in water used for handwashing. *bla*NDM was present in the HWW of one Beninese  
36 hospital and was also detected in the stools of a hospitalized patient. Mobile colistin resistance genes (*mcr*) were  
37 detected in the HWWs of all the three studied countries, with *mcr*-5 variants being the most common. These and  
38 other *mcr* genes were observed in very high abundance in treated wastewater released into rivers in Burkina Faso.  
39 In Benin, only little is done to treat wastewaters, including HWWs. The results provide evidence for public health

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

40 decision-makers in Benin and Burkina Faso for the dire need to increase wastewater treatment capacity, with  
41 particular attention to HWWs.

42

43 *Keywords*

44 antimicrobial resistance (AMR)

45 West Africa

46 hospital wastewater (HWW)

47 carbapenemase

48 colistin resistance

49 metagenomes

50

51 1. Introduction

52

53 The AMR crisis is a global issue with multifaceted effects on human, animal, and environmental health, and it  
54 comes with substantial economic losses (1). Resistance to almost all antibiotics, including the last resort regimen  
55 reserved to treat the most severe infections, has tended to emerge only a couple of years after introducing a new  
56 antibiotic to the global market (2, 3). Due to limitations in diagnostic testing in low-resource settings in low-  
57 income countries (LICs), broad-spectrum antibiotics are often prescribed empirically without microbiological  
58 verification of the causative pathogen or its sensitivity to different antibiotics (4). In addition, unregulated access  
59 to antimicrobials allows their wide use in these countries (5–7).

60

61 Bacterial resistance to antimicrobial compounds may be intrinsic or acquired from other bacteria or the  
62 surrounding environment via horizontal gene transfer of mobile genetic elements (MGE) (8–10). Acquired,  
63 potentially mobile ARGs are considered to have greater clinical relevance and impact the current AMR problem  
64 induced by human activity (10, 11). Class 1 integrons are strongly linked to the dissemination of clinically relevant  
65 acquired ARGs (12, 13). Although integrons are not mobile themselves, integron-carried multidrug-resistance  
66 gene cassettes can be transmitted to new hosts, for example, via plasmids (14). The gene encoding the integron  
67 integrase (*intI*), among other associated genes, is involved in the recombination of incoming genetic material into  
68 the gene cassette, and *intII* is typically used as a marker for class 1 integrons (13, 15). Over time, quaternary  
69 ammonium compound (QAC) resistance genes, such as *qacEΔ*, have established their position in the structure of  
70 a typical class 1 integron gene cassette whose spread is induced by human activity (15, 16). It has been speculated  
71 that increasing the use of such disinfectant compounds in hospitals and community environments, especially  
72 during the COVID-19 pandemic, would further exacerbate the problem of multidrug resistance mediated by class  
73 1 integrons (17).

74

75 The high-risk ESKAPEc bacteria (9) (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*,  
76 *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, *Enterobacter* spp., and *Escherichia coli*) (18) are top  
77 priority pathogens for AMR research and the development of new antibiotics (9, 19). They can feature an  
78 impressive resistance profile in terms of both intrinsic and acquired ARGs and often exhibit multidrug-resistant  
79 phenotypes. However, investigating only bacterial isolates may result in a limited understanding of ARGs

80 occurring in a particular hospital. In contrast, shotgun metagenomic sequencing enables the analysis of resistomes  
81 and mobilomes (all ARGs or MGEs, respectively) as well as the microbial communities present in the studied  
82 environment (20). Hence, as opposed to PCR-based methods, there is no requirement for a priori knowledge of  
83 the resistome (20).

84

85 As a consequence of the increased prevalence of extended-spectrum beta-lactamase-producing Enterobacteriaceae  
86 (ESBL-PE), the use of broad-spectrum antibiotics has increased in clinical practice (2, 21, 22). Tracking down  
87 the spread of ESBL genes is, in many situations, no longer feasible or relevant, and attention has been directed to  
88 plasmid-mediated genes that encode carbapenemases. The carbapenemases *bla*GES, *bla*IMP, *bla*KPC, *bla*NDM,  
89 *bla*OXA-48, *bla*OXA-58, and *bla*VIM carried by plasmids have emerged and spread around the world during the  
90 past three decades (23–26).

91

92 Metronidazole is an example of an empirically widely used first-line treatment in many African countries (27).  
93 This antimicrobial drug is effective against various protozoan parasites and anaerobic bacteria (28). Although  
94 acquired resistance towards metronidazole (mediated mainly by *nim* genes) is considered to be relatively rare, an  
95 increasing trend in its emergence has been reported (28). Colistin is a last-resort antibiotic used for treating  
96 infections caused by multidrug and extensively resistant bacteria (29). The rapid emergence of mobile colistin  
97 resistance mediated by *mcr* genes threatens the efficacy of colistin in clinical use (30, 31).

98

99 Although AMR is of global concern, the crisis affects most dramatically LICs, such as those in West Africa (3, 4,  
100 32–34). Lack of research data is a major factor hindering the development of solutions to tackle the AMR problem  
101 in these countries (4, 32, 35, 36). Despite gaps in resistance surveillance data, it is well known that the level of  
102 AMR is highest in many LICs, such as in many African countries (21, 33, 37, 38). By contrast, in Northern  
103 European countries, such as Finland, AMR occurrence is among the lowest globally, both in the community (38,  
104 39) and healthcare settings (40).

105

106 As hospital wastewater (HWW) from healthcare facilities is considered to be at the frontline of AMR emergence  
107 and spread (41), we set out to investigate the AMR situation and the characteristics of the resistomes in nine  
108 hospitals in two West African countries, Benin and Burkina Faso, where prior data was scarce (21). For  
109 comparison, we analyzed samples from four hospitals in Finland, where the level of AMR was expected to be low  
110 (40). In this study, we applied shotgun metagenomic sequencing as it allows the analysis of the total catalog of  
111 genes present in the studied environment.

112

## 113 2. Results

114

### 115 2.1. Features of hospital wastewaters in Benin, Burkina Faso, and Finland

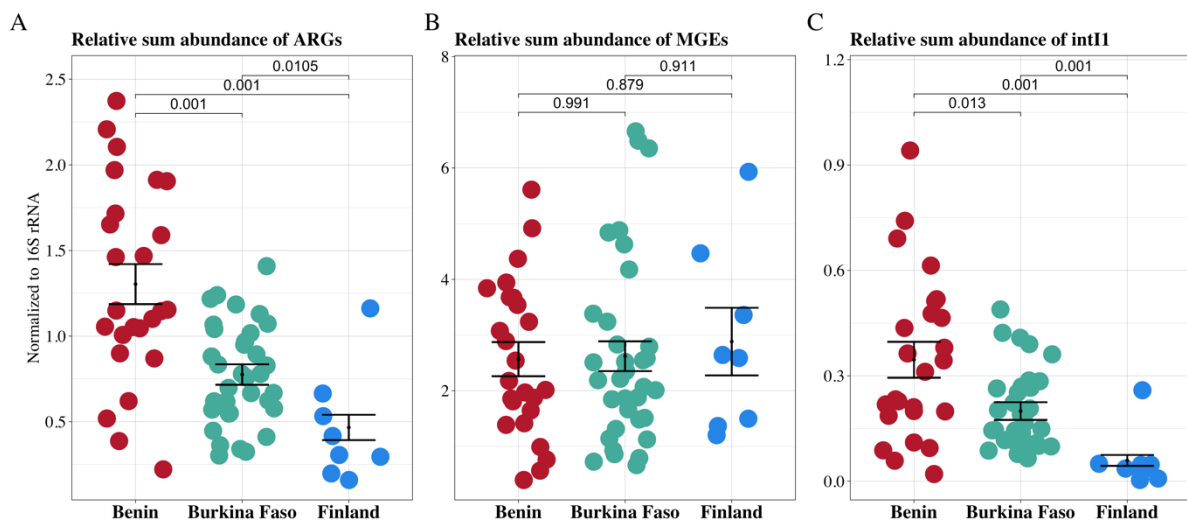
116

117 We analyzed hospital wastewater-associated resistomes from Benin, Burkina Faso, and Finland. On average, 31  
118 million sequence reads per sample were analyzed. HWWs from Benin showed the highest, and from Burkina Faso  
119 the second-highest abundance of all detected ARGs normalized to bacterial 16S rRNA genes, here referred to as

120 the relative sum abundance of ARGs. The lowest relative sum abundance was detected in HWWs from Finland  
121 (Figure 1A, Table S3A). No significant differences were observed in the relative ARG sums among the different  
122 hospitals in each country or between the hospital sections or wards within the sampled hospitals (Table S3A).  
123 Country-wise differences for the relative abundances of class 1 integron gene (*intI1*) (Figure 1C, Table S3C)  
124 followed the trend observed for relative ARG abundances (Figure 1A). The observed high Pearson correlation of  
125 0.80 ( $p = 3.2 \times 10^{-16}$ ) between ARGs and *intI1* genes across all 67 HWW samples indicated an association between  
126 the genes. For all MGEs, including also other than class 1 integron associated genes, a similar trend was not  
127 observed (Figure 1B, Table S3B).

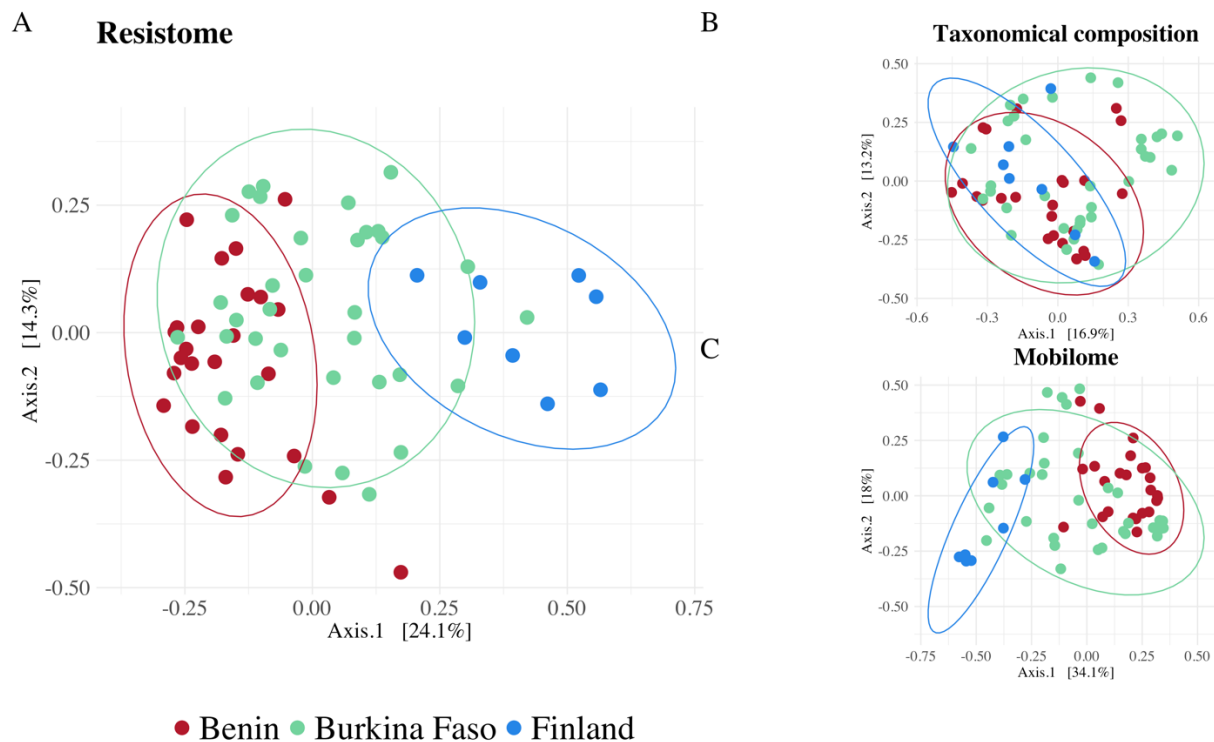
128  
129 The resistomes of HWWs from Benin and Burkina Faso resembled each other more than those of the Finnish  
130 HWWs (Figure 2A). The resistomes in Benin and Burkina Faso shared 384 ARGs, while Finland and Benin shared  
131 only 15 ARGs, and Finland and Burkina Faso shared 134 ARGs (Figure S4). The resistome compositions differed  
132 the most between the Beninese and Finnish HWW samples (Figure 2A). The country explained the differences in  
133 the resistomes by 29.5–33.7 % in the Benin-Finland comparison and 24.5–25.6 % in the Burkina Faso-Finland  
134 comparison (Table S1). In addition, while the ARGs found uniquely in HWWs from Burkina Faso made up  
135 approximately 14 % of all the detected ARGs in this study, the corresponding percentage for HWWs from Benin  
136 was only 2.5 % (Figure S4).

137  
138 As the resistomes of West African countries and Finland were divergent, the drivers for the differences were  
139 subsequently investigated. The country explained less of the variance between taxonomical compositions than  
140 resistomes (Figure 2B, Table S1). A set of samples from a single hospital (Hospital I, Figure S3B) seemed to drive  
141 the distinct clustering seen in the microbial community in HWWs from Burkina Faso.  
142



143  
144  
145 **Figure 1.** Relative sum abundance of ARGs, MGEs, and class 1 integrons. A generalized linear model (GLM)  
146 was fitted to 16S rRNA, and gene length normalized A) ARG B) MGE, and C) *intI1* count data to model the  
147 differences in the levels of these genes in the three countries. The vertical line shows the median in boxplots,  
148 while the lower and higher hinges denote the 25 % and 75 % quantiles, respectively. In addition, Wilcoxon signed-  
149 rank test significances are shown for pair-wise comparisons.

150



151

152 **Figure 2.** PCoA plots describe the ordinations based on A) resistomes, B) taxonomical compositions, and C)  
153 mobilomes in HWWs from Benin, Burkina Faso, and Finland. All 67 HWW samples were included in the plots.  
154 However, due to unequal group sizes, groups of random sample sets were applied for testing the significance of  
155 the clustering based on the country using adonis. The tests were performed in triplicate with different samples to  
156 account for variation caused by sample selection. The  $R^2$  values are shown in Supporting Information, Table S1.  
157 Confidence ellipses represent 90 % confidence levels.

158

## 159 2.2. Clinically relevant taxa and associated ARGs

160

161 Taxonomical composition present in septic tanks and sumps in Benin and Burkina Faso differed from those found  
162 in hospital sewers in Finland. In Benin and Burkina Faso, the 15 most abundant genera present in HWW included  
163 *Desulfobulbus*, *Sulfurospirillum*, *Thauera*, and *Geobacter*, while in Finnish HWW, *Bifidobacterium* and  
164 *Faecalibacterium* were among the most abundant genera (Table S4). Genera known for opportunistically  
165 pathogenic species such as *Escherichia* and *Aeromonas* were detected among the 15 most abundant genera in the  
166 HWWs of all three countries (Table S4, Figure S5). *Klebsiella pneumoniae* was among the 15 most abundant  
167 species in Benin and Burkina Faso in terms of relative abundance (Table S4, Figure S5). In contrast, at the species  
168 level, *Escherichia coli*, *Enterococcus faecium*, and species of *Enterobacter cloacae* and *Pseudomonas aeruginosa*  
169 complexes were present but rare in all the countries. *Staphylococcus aureus* was detected exclusively in Burkinabe  
170 samples (Figure S5). Other species in the genus *Staphylococcus* were more commonly detected in HWWs from  
171 Benin, and Burkina Faso compared to Finland (Tables S5A and S6A). However, knowing the limitations of  
172 species-level taxa identification using shotgun metagenomics (42), these results must be evaluated with some  
173 caution.

174

175 There was no significant difference (adjusted  $p$ -value  $< 0.05$ ) in the abundance of the genus *Acinetobacter* between  
176 the HWWs from the three countries, but the abundances of different species of *Acinetobacter* varied significantly.  
177 *A. johnsonii* and carbapenemases in the *bla*OXA-211-like group (for example, *bla*OXA-211, -212, -280, and -281)  
178 that are intrinsically carried by this species were significantly more abundant in HWWs from Finland compared  
179 to either Benin or Burkina Faso (Figure 4, Figure S5 and Tables S5A, S6A, S7, and S8). Certain species of  
180 *Acinetobacter* that are not intrinsically carbapenem-resistant, such as *A. townerii*, were detected at significantly  
181 higher relative abundances in HWWs from Benin and Burkina Faso compared to Finland (Tables S5A, and S6A).

182

### 183 2.3. Carbapenem resistance

184

185 The presence of seven acquired carbapenemases (*bla*GES, *bla*IMP, *bla*KPC, *bla*NDM, *bla*OXA-48, *bla*OXA-58,  
186 and *bla*VIM) was analyzed in detail due to their ability to cause complex infections with minimal treatment options  
187 when in a relevant host. One of the Finnish (hospital K) samples possessed the highest relative abundance of these  
188 carbapenemases (Figure 3A). This hospital is the University Central Hospital of the capital of Finland. In HWWs  
189 from Benin, the carbapenemase genes in the *bla*GES family seemed to dominate over other carbapenemases in  
190 all four hospitals (Figure 3A). Instead, most of the other carbapenemases were present at similar prevalence levels  
191 in HWWs from Burkina Faso and Finland (Figure 3A), in line with the observation of the lower number of  
192 different ARGs in Benin versus Burkina Faso (Figure S4).

193

194 *bla*VIM, *bla*IMP, and *bla*NDM were detected in Burkinabe and Finnish HWW samples, but they were more  
195 abundant in HWWs from Burkina Faso (Figure 3A). Among the studied acquired oxacillinases, *bla*OXA-48 was  
196 detected solely in two samples from Burkina Faso (Figure 3A). Instead, *bla*OXA-58 was present in the majority  
197 of Burkinabe and Finnish samples (Figure 3A). Interestingly, most *bla*KPC genes were exclusively detected in  
198 HWWs from Finland (Figure 3A, Table S11).



199  
 200 **Figure 3.** The abundance of carbapenemase genes in relation to 16S rRNA genes in A) HWWs from Benin,  
 201 Burkina Faso, and Finland, and B) various other sample types in Benin and Burkina Faso. The hospital code  
 202 (Figure 3A) and sample names (Figure 3B) are marked on the x-axis. Only the *blaGES* genes that encode  
 203 carbapenemase production (*blaGES*-2, -4, -5, -6, -14, -16, -20 (25)) according to the Beta-lactamase Database  
 204 (43) are shown, omitting the *blaGES* genes that encode ESBL. All screened carbapenemase gene variants are  
 205 listed in Supporting Information, Table S18.

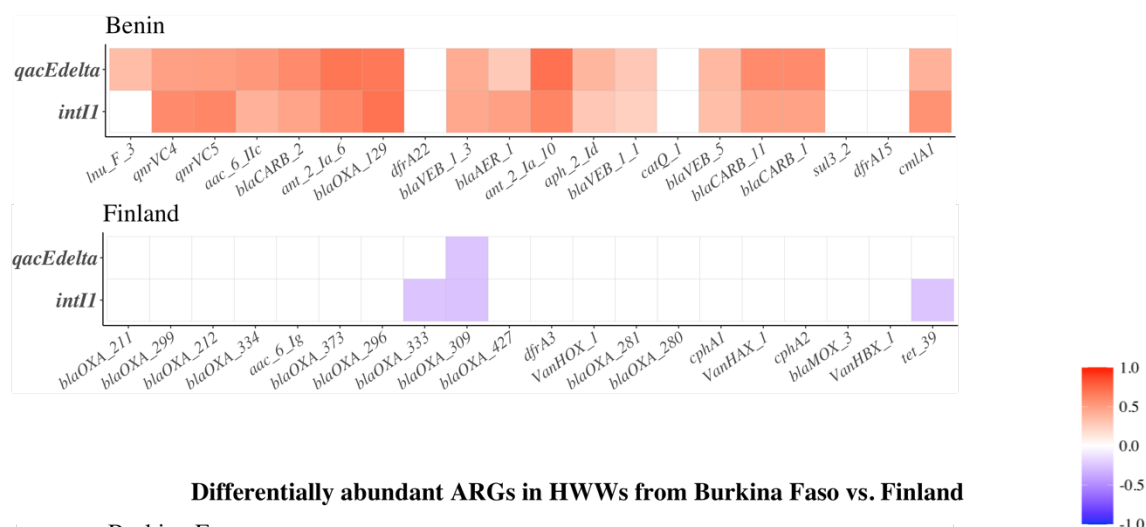
206  
 207 2.4. Putative role of MGEs and class 1 integrons in shaping HWW resistomes

208  
 209 Three methods were used to study whether the abundance of class 1 integrons could explain the differences in the  
 210 relative sum abundances of ARGs between the HWWs of the three countries (as shown in Figure 1A). First,  
 211 correlations between the relative abundances of the detected ARGs and the two class 1 integron-associated genes,  
 212 *intI1* and *qacEA*, were studied (Tables S12A and S12B). These correlations were visualized for those ARGs  
 213 differentially abundant in HWWs from each country in pair-wise comparisons (Figure 4). Second, contigs carrying  
 214 the *intI1* gene were studied for the putative co-localization of ARG sequences (Table S13). Third, fragments  
 215 carrying multiple acquired ARGs were screened visually using assembly graphs (Figure S6).

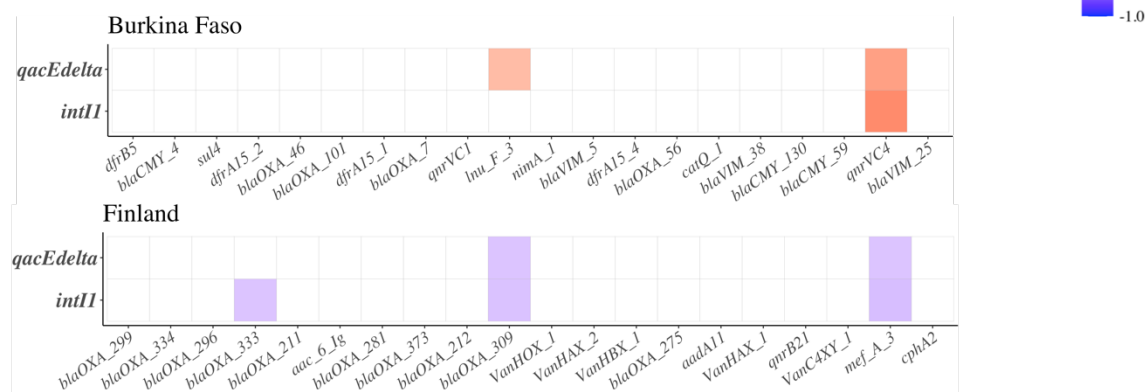
216  
 217 The top 20 ARGs that were more abundant in the Beninese HWWs showed stronger positive correlations with  
 218 class 1 integron genes compared to the ARGs that were more abundant in Burkina Faso and Finland (Figure 4).

219 This suggests a stronger class I integron association of ARGs in Beninese HWWs than in the other two countries.  
 220 The detection of some of these ARGs in contigs carrying several other acquired ARGs (Figure S6) or class 1  
 221 integron-associated genes (Table S13) suggests that the genes are linked to class 1 integrons. For example, the  
 222 plasmid-mediated ARGs *qnrVC4* and *qnrVC5* strongly correlated with *intI1* and *qacEΔ*. The genes were located  
 223 on contigs with other ARGs such as aminoglycoside ARGs and either carbapenemase (Figure S6) or ESBL genes  
 224 in the *blaGES* gene family (Figure S6B). While *qnrVC4* and *qnrVC5* were more abundant in HWWs from Benin  
 225 and Burkina Faso than in those from Finland, quinolone ARGs unassociated with integrons, such as *qnrB* variants,  
 226 were instead more common in HWWs from Finland (Figure 4, and Tables S7, and S8).  
 227  
 228 Oxacillinases constituted half of the top 20 most abundant ARGs in HWWs from Finland (Figure 4). Most of  
 229 these oxacillinases (*blaOXA*-211, -212, -334, -373, -309, -280, and -281) belong to the group of *blaOXA*-211-  
 230 like carbapenemases typically intrinsically carried by certain *Acinetobacter* species and showed no significant  
 231 correlation with *intI1* or *qacEΔ* (Figure 4, Tables S12A and S12B). In contrast, certain genes of oxacillinases  
 232 from other *blaOXA* gene families, which were significantly more abundant in HWWs from Benin or Burkina  
 233 Faso, showed significant correlation with *intI1* and *qacEΔ* (Figure 4, Tables S12A and S12B) and were found in  
 234 shared contigs with the *intI1* gene and other ARGs (Table S13, Figure S6C).

**Differentially abundant ARGs in HWWs from Benin vs. Finland**



**Differentially abundant ARGs in HWWs from Burkina Faso vs. Finland**



235  
236



237 **Figure 4.** Differently abundant ARGs in HWWs from Benin and Burkina Faso and their putative associations  
238 with class 1 integrons. 20 significantly more abundant ARGs in HWWs of each country in pair-wise comparisons  
239 (Benin-Finland, Burkina Faso-Finland) are shown. Pearson correlations between each ARG and the genes *intI1*  
240 and *qacEA* associated with class 1 integrons (Supporting Information, Tables S12A, and S12B) are denoted with  
241 color-coding. The red color indicates a positive and blue negative correlation. Non-significant ( $p > 0.05$ )  
242 correlations are marked as blank (white) squares.

243

#### 244 2.5. Mobile colistin resistance genes

245

246 Mobile colistin resistance genes (*mcr* genes) were observed in several HWW samples in Benin, Burkina Faso,  
247 and Finland (Table S16). Variants of *mcr-5* (*mcr-5.1* and *mcr-5.2*) were the most common as they were found in  
248 HWWs from all except one hospital (hospital F) in the three countries. We were able to confirm the location of  
249 *mcr-5* within a Tn3-like transposase element in the metagenome-assembled contigs (Figure S7, Table S14).

250

251 In addition to the HWW samples, high relative abundances of *mcr-5* genes were also detected in the immediate  
252 and more distant surroundings of the hospitals, such as in the street gutter water 100 m away from hospital B and  
253 in the soil between two HWW septic tanks in Benin (Table S15A). For hospital C, the relative abundance of *mcr-*  
254 *5* in a puddle surrounding the surgery room HWW septic tank was greater than the average relative sum abundance  
255 of *mcr-5* in the actual HWW septic tanks of that hospital (Figure S2D, Table S15A).

256

257 In Burkina Faso, very high relative sum abundances ( $2.8 \times 10^{-2}$ ) of *mcr-5* were observed in a municipality channel  
258 leading out from hospital I after passing through a biological treatment process (Table S15B). Furthermore, the  
259 presence of *mcr-5* was also detected in a nearby wetland where the processed wastewater is released (Table S15B).  
260 Similarly, in another Burkinabe city, *mcr-5* was found in a sample before wastewater treatment and the river  
261 receiving the water after treatment (Table S15B).

262

263 The next most prevalent *mcr* genes, namely *mcr-3*, *mcr-10*, and *mcr-7*, were detected in certain cases at very high  
264 relative abundances in the HWW samples (Table S16). The gene *mcr-3* was also present in the water used for  
265 handwashing at hospital C (Figure S2E, Table S15A). The gene *mcr-2* was not detected, and the lowest possible  
266 average relative sum abundance was detected for the gene *mcr-1* (Table S16).

267

#### 268 2.6. Other than hospital wastewater samples

269

270 Other sample types within and outside the studied hospitals were investigated to obtain a more comprehensive  
271 understanding of ARG prevalence. Metronidazole resistance genes, such as *nimA* and *nimE*, were among the top  
272 15 most abundant ARGs in the fecal samples of the four in-patients at a Beninese hospital B (Table S17). These  
273 ARGs were also significantly more abundant in HWWs from Benin, and Burkina Faso compared to Finland  
274 (Figure 4, Tables S7 and S8). *bla*NDM was detected in one of the four fecal samples (Figure 3B) and was also  
275 present in the HWWs of hospital B (Figure 3A). Other high abundance ARGs in the fecal samples included ESBL  
276 and broad-spectrum beta-lactamase genes of the *cfxA* family (Table S17).

277

278 As in the HWWs from Benin (Figure 3A), *bla*GES carbapenemases were also prevalent in other sample types  
279 from a hospital environment, such as in the water intended for handwashing (Figure S2E) for hospital C, the soil  
280 between HWW septic tanks in the yard of hospital B, and street gutter water 100 meters away from hospital B  
281 (Figure 3B). *bla*GES carbapenemase genes were also found in various waters used for drinking, including river  
282 waters in a remote countryside village (Figure 3B). However, the relative sum abundance of ARGs was low in  
283 the three rivers in central Benin (Figure 3B). We also detected *bla*NDM in well water used for drinking in Benin  
284 (Figure 3B). Albeit at low relative abundances, *bla*GES, *bla*VIM, and *bla*OXA-58-like carbapenemases were  
285 detected in natural waters receiving treated HWWs from Burkina Faso (Figure 3B). Significantly higher  
286 abundances of these ARGs were found in biologically treated wastewater before releasing into the river (Figure  
287 3B).

288

### 289 3. Discussion

290

291 We characterized the bacterial community composition and resistome in 67 HWWs and 16 related samples from  
292 Benin and Burkina Faso and compared them to HWWs from Finland. This is the first study investigating hospital  
293 wastewaters in the two West African countries using a metagenomic approach. Due to the lack of systematic AMR  
294 surveillance in these countries, available data on bacterial resistance are patchy and heterogeneous (1, 4, 31, 34,  
295 37). Thus, the magnitude of the resistance problem and the specific ARG reservoirs in many African countries are  
296 yet to be unraveled. The relative sum abundance of all ARGs detected in HWWs was the highest in Benin, yet the  
297 number of different ARGs detected there was the lowest. The differences in resistomes between the countries  
298 seemed to be only partially explained by variation in taxonomical composition, as the clustering of taxonomical  
299 composition was not as strongly correlated by country compared to clustering of resistomes. This suggests that  
300 despite differences in the HWW collection systems in the studied countries (HWW is relatively stagnant in septic  
301 tanks or sumps in Benin and Burkina Faso and flowing in HWW drains in Finland), the microbial community  
302 compositions appear to be similar. However, among the most abundant taxa, we did find many environmental  
303 bacteria in HWWs from Benin and Burkina Faso but not in Finland. This may be due to the closer contact of water  
304 with environmental surfaces in septic tanks and sumps compared to the more isolated wastewater systems in  
305 Finland. Bacterial genera typically found in soil and aquatic environments (44) have also been previously  
306 described to dominate HWWs from Benin and Burkina Faso (45). In contrast, the 15 most abundant genera in  
307 Finnish HWWs included less environmental bacteria and more bacteria typical for human gut communities, such  
308 as *Bifidobacterium* (46) and *Faecalibacterium* (47). Another factor possibly explaining the differences in the  
309 dominant taxa present in the HWWs from Benin and Burkina Faso in comparison to Finland could arise from the  
310 fact that in Finland, the hospital toilet waters containing human fecal material are directed to HWWs, while for  
311 the majority of the West African HWWs studied here, this was not the case.

312

313 The ARGs that were more abundant in Benin's HWWs than in Finland appeared to be more commonly carried by  
314 class 1 integrons. They were strongly correlated with class 1 integron-associated genes (*intI1* and *qacEA* (15))  
315 and, in certain cases, detected in shared contigs with these genes. Moreover, we could detect gene cassettes  
316 consisting of several different acquired ARGs linked to each other, likely referring to gene cassettes carried by

317 class 1 integrons (48). Furthermore, a higher abundance of ARGs intrinsically carried by specific species, such as  
318 *blaOXA-211* and *blaOXA-299* genes by *A. johnsonii* and *A. bouvetii* (49, 50), as well as *cphA* intrinsically carried  
319 by *Aeromonas* spp. (51) were distinctive features of the HWW resistomes from Finland. Intrinsic resistance is  
320 often chromosomally encoded and rarely mobile or related to horizontal gene transfer (52), whereas acquired  
321 resistance is typically carried by MGEs, such as class 1 integrons. As expected, no significant correlations between  
322 these intrinsic *blaOXA* genes and class 1 integron-associated genes were found, supporting the observation that  
323 ARGs putatively carried by class 1 integrons had a smaller impact on resistomes in HWWs from Finland compared  
324 to Benin. Based on these results, we speculate that the high sum abundance of ARGs seen in HWWs from Benin  
325 may, in part, be explained by the mobilome, more specifically, the prevalence of class 1 integrons. However,  
326 correlation does not necessarily indicate causation, and in this study, quantitative inferences cannot be made for  
327 metagenome-assembled contigs and the putative co-localization of ARGs and class 1 integron-associated genes.  
328 MGEs other than class 1 integron-associated genes did not exhibit a significant correlation with ARGs. This is in  
329 line with the current understanding of the role of class 1 integrons, in particular, among all MGEs in spreading  
330 acquired ARGs (13).

331  
332 Interestingly, the ARGs observed in HWWs from Benin were the highest in number but probably lower in clinical  
333 importance compared to those in the HWWs from Burkina Faso and Finland. While the most significant  
334 carbapenemase genes *blaKPC*, *blaNDM*, *blaVIM*, *blaIMP*, *blaOXA-48*, and *blaOXA-58* were found in several  
335 HWW samples from Burkina Faso and Finland at varying abundances, the abundance of many of them was low,  
336 or they were not detected at all in HWWs from Benin. The *blaGES* carbapenemase genes were an exception. They  
337 were the most prevalent of the seven carbapenemases in Beninese HWWs and were also found in samples  
338 unrelated to hospitals.

339  
340 We suspect that as *blaGES* carbapenemases are typically carried by class 1 integrons, as confirmed by previous  
341 studies (53), a co-selection phenomenon (3, 54) might play a role in increasing their spread also in here. In this  
342 scenario, residues of antibiotics other than carbapenems would cause the co-selection of *blaGES*. We also found  
343 a high relative abundance of plasmid-mediated quinolone ARGs in the HWWs from Benin. These *qnrVC* genes  
344 showed strong correlations to class 1 integron-associated genes and were in some cases found to be located in the  
345 same contigs with *blaGES-5* carbapenemase genes (Figure S6B). Hence, selection targeting quinolone ARGs  
346 could co-select for carbapenem resistance encoded by the gene *blaGES-5*. Altogether our observations of ARGs  
347 putatively carried by class 1 integrons are worrisome as new ARG combinations among integron gene cassettes  
348 enable their selection by a broader set of antimicrobials or biocides such as quaternary ammonium compounds  
349 (QACs) (15, 55, 56).

350  
351 Despite the lowest relative sum abundance of ARGs detected in HWWs from Finland, some of the seven  
352 carbapenemase genes of particular interest seemed to have even higher abundances in Finland than in the surveyed  
353 West African countries. However, the occurrence of these ARGs was not homogenous among the Finnish samples.  
354 The highest relative abundance seen in especially one sample from the University Central Hospital of the capital  
355 of Finland (hospital K) was likely explained by the high patient turnover compared to other hospitals in Finland  
356 and Benin and Burkina Faso studied here. Furthermore, a high relative abundance *blaKPC* was observed in only

357 one of the Finnish samples. Furthermore, the overall occurrence of *blaKPC* across all HWWs from Benin, Burkina  
358 Faso, and Finland studied here aligned with many local and global reports and systematic reviews (57–60),  
359 according to which *blaKPC* carbapenemases are spreading more profusely in Europe and North America than in  
360 Africa. Similarly, the relative abundance of *blaOXA-58* was higher among Finnish than Beninese or Burkinabe  
361 HWW samples.

362  
363 Disconcertingly, *blaNDM* was found in the fecal sample of an in-patient staying at a Beninese hospital. *blaNDM*  
364 was also detected in the wastewaters of the hospital, indicating the sensitivity of wastewater-based monitoring of  
365 bacterial resistance carried by hospital in-patients. *blaNDM* carried by *A. baumannii* has recently, in 2021, been  
366 reported in clinical settings in Benin (61).

367  
368 Our observation of the high prevalence of metronidazole ARGs in Benin and Burkina Faso but not in Finland  
369 most likely reflects differences in the consumption of this drug. In LICs, diarrheal infections caused by pathogens  
370 such as *Clostridioides difficile* and other anaerobic bacteria, as well as protozoa, are common. Therefore, the  
371 consumption of metronidazole commonly used to treat these infections may be high in these countries as these  
372 drugs are being prescribed, for example, both with and without a positive diagnostic test for *C. difficile* (62).  
373 However, multidrug-resistant strains of *C. difficile* seem to already be prevalent in hospitals in certain parts of  
374 Africa as a consequence of the heavy use of antimicrobials among in-patients(62).

375  
376 Compared to many HMICs such as those in Northern Europe, antimicrobial usage in agricultural (6, 7) and clinical  
377 (4) settings are less controlled – or even entirely unregulated – in many African countries. For example, while  
378 banned in many other countries, the use of colistin as a feed additive is allowed in many LICs (63), such as those  
379 in Africa (7). In contrast to previous reports on the prevalence of mobile colistin resistance globally (30) and in  
380 Africa (31, 64), *mcr-5* was the most commonly detected gene in the HWWs from our study. Instead, our  
381 observations are in line with the results obtained by using both metagenomics and qPCR (65). This discrepancy  
382 may be due to differences between the methodologies used to screen for colistin resistance. Colistin ARGs other  
383 than the best-known genes *mcr-1*, *-2*, and *-3* are rarely targeted when screening for mobile colistin resistance with  
384 conventional PCR (65). Our study shows that to obtain a more realistic view of mobile colistin resistance genes  
385 in Africa, screening should also be conducted for different *mcr* genes, as was recently done by Ngbede and  
386 colleagues (66).

387  
388 The *mcr-5* gene detected here was embedded in a Tn3-like transposase similar to previous reports for *Salmonella*  
389 *enterica* (67, 68) and *Escherichia coli* (69) plasmids and the chromosome of *Cupriavidus gilardii* (67). These  
390 Tn3-like elements are flanked by inverted repeats, which enable translocation and putatively a wide host range  
391 for the *mcr-5.1* harboring element (67–69). Based on our results, the occurrences of *mcr-5* and other mobile  
392 colistin resistance genes in Benin and Burkina Faso were not restricted to HWW septic tanks. For example, at a  
393 Beninese hospital, a water puddle in the hospital yard had a higher relative abundance of *mcr* genes compared to  
394 the actual HWW septic tanks for that hospital. Furthermore, these genes were also detected in water intended for  
395 handwashing of visitors to the hospital. Thus, the transmission of these clinically critical ARGs within hospitals  
396 could be enabled via hands contaminated by this water. Similarly, Jacobs and colleagues have drawn attention to

397 the potentially inferior microbiological water quality in similar hand washing water tanks, which are very common  
398 in West Africa (4).

399

400 In Burkina Faso, the relative abundance of *mcr-5* released into natural waters after wastewater treatment was high  
401 ( $2.36 \times 10^{-4}$ ). For comparison, a similar relative abundance of *mcr-5* genes ( $3.0 \times 10^{-4}$ ) was reported for untreated  
402 municipal wastewater in Germany (65). Although we cannot confirm the association between the *mcr-5* genes  
403 detected in the natural waters and the hospital-associated wastewater in the Burkinabe hospital, our study suggests  
404 the inability of the currently used wastewater treatment processes in Burkina Faso to remove colistin resistance  
405 genes. Wastewater treatment systems have also earlier been described to be inadequate in Benin and Burkina Faso  
406 (70). This increases the risk of the further spread of colistin and other ARGs as the inadequately treated water is  
407 released into natural waters and used for various purposes by local people. Furthermore, in many African  
408 countries, untreated wastewater is used for irrigation in urban agriculture, possibly enabling the dissemination of  
409 ARGs to fresh produce (71, 72). The situation appears to be especially critical in Cotonou, where the hospitals  
410 involved in the study did not use any kind of wastewater treatment. Groundwater in the city is extracted from the  
411 shallow aquifer, which is polluted due to unauthorized waste deposits, inadequate toilets, pit latrines, and septic  
412 tanks prone to leakage and hydraulic failure (73). Our study shows that untreated hospital wastewaters, which can  
413 leak into groundwater and the environment, carry clinically dangerous antimicrobial-resistant bacteria and  
414 resistance genes.

415

416 Taken together, based on our results from nine hospitals, we can state that clinically important ARGs are  
417 circulating in West African hospitals and can spread to their surroundings. There was no significant difference in  
418 the relative sum abundances of ARGs based on the hospital section or different hospitals within these countries.  
419 Various carbapenem ARGs were found in HWWs from Burkina Faso and Finland, while among these *bla*GES  
420 seem to dominate in the studied hospitals in Benin. Our results suggest that class 1 integron mediated co-selection  
421 could play an important role in shaping the resistome in the Beninese samples. In addition to HWWs,  
422 carbapenemase genes and mobile colistin resistance genes were detected in the vicinity as well as locations more  
423 distant from the hospitals. Finally, there were differences between the HWW resistomes in Benin and Burkina  
424 Faso concerning the relative sum abundance of ARGs and the prevalence of clinically relevant ARGs. Therefore,  
425 it is important to consider the differences between African countries in AMR surveillance.

426

## 427 4. Material and Methods

428

### 429 4.1. Sample description

430

431 Hospital wastewater (HWW) samples were collected in Benin from four (hospitals A-D) and in Burkina Faso  
432 from five different hospitals (hospitals F-J) in November and December of 2019. In Finland, HWW samples were  
433 collected from four different hospitals (hospitals K-N) in January 2020. In Burkina Faso, the samples were mainly  
434 from septic tanks or sewers in the hospital area. In Benin, none of the hospitals was connected to a sewer system,  
435 and the samples were from septic tanks or sumps (unstructured wastewater wells), which were never emptied to  
436 our knowledge. In most cases, the toilet water was not directed into these sumps. For comparison, some other

437 types of samples were included: samples from river waters used for drinking in a remote countryside village in  
438 the community of Savalou in central Benin (n=3); river or wetland water receiving treated wastewater in Burkina  
439 Faso (n=2, hospitals I and J); and fecal samples of in-patients in Benin (n=4, hospital B). In addition, the following  
440 samples were analyzed; well water (hospital A), treated tap water (community of Savalou in central Benin), water  
441 for handwashing (hospital C), water in a municipality channel receiving biologically treated wastewater (hospital  
442 I), water and sediment from a street gutter next to the hospital (hospital B), an empty HWW septic tank (hospital  
443 A), and the soil between HWW septic tanks (hospital B). The four fecal samples were obtained from patients  
444 staying at hospital B with diagnoses mainly associated with musculoskeletal injuries. No clinical data concerning  
445 bacterial infections were reported for these patients, and we have no information on their possible use of  
446 antibiotics. Patient age varied from 24 to 50 years old, and at the time of sampling, patients had stayed in the  
447 hospital for one to two months. All sample details are provided in the Supporting Information, Table S2A. Selected  
448 pictures and a map indicating sample collection regions are shown in Figures S1 and S2. Letters A–N were used  
449 to denote the hospitals/sampling locations.

450

#### 451 4.2. DNA extraction and metagenomic sequencing

452

453 Water samples were collected into 1-liter bottles and transported to a laboratory on ice. They were kept at +4 °C  
454 until processed within 24 hours. A volume of 50–100 ml was filtered through a 0.2 µm polycarbonate filter  
455 (Whatman™, GE Healthcare Life Sciences) using a portable vacuum pump (Millivac-Mini Vacuum Pump XF54,  
456 Millipore, Merck). DNA was extracted from the filters using the Qiagen DNeasy PowerWater DNA Kit for the  
457 water samples and Qiagen DNeasy PowerSoil DNA Kit for the soil samples, following manufacturer instructions.  
458 The concentration and quality of the extracted DNA were determined with a NanoDrop spectrophotometer.  
459 Altogether 93 samples were subjected to shotgun metagenomic sequencing using Illumina Novaseq5000 with  
460 Nextera XT library preparation at the Institute of Biotechnology, University of Helsinki.

461

#### 462 4.3. Bioinformatic analyses

463

464 All samples were analyzed using an in-house Snakemake workflow (74). Briefly, the quality control steps  
465 included in the workflow were performed using FastQC (75) and MultiQC (76), with adapter and low-quality  
466 read removal using Cutadapt (77). Nucleotide sequence reads were mapped using Bowtie2 (78) against the  
467 ResFinder database (version 3.2, downloaded on 2020-06-28) (79). The reads were sorted and filtered using  
468 SAMtools (80) such that the reads mapping as pair or alone were calculated as a single count. Mobile genetic  
469 elements were identified by mapping the reads against the MobileGeneticElementDatabase  
470 (<https://github.com/KatariinaParnanen/MobileGeneticElementDatabase>, downloaded on 2020-06-28) consisting  
471 of 2714 unique MGE sequences including transposases; integrons of classes 1, 2, and 3; the integron associated  
472 disinfectant resistance gene *qacEΔ*; and IS and ISCR type insertion sequences (46). Class 1 integrons were studied  
473 using mapping results obtained for the *intI* gene, which encodes an integron integrase (*IntI*). Taxonomic profiling  
474 was performed using both Metaphlan3 (81) and Metaxa2 (82). The counts for bacterial 16S rRNA from Metaxa2  
475 were used to normalize ARG and MGE counts. The lengths of the functional genes and the 16S rRNA SSU gene  
476 were taken into account in the normalization.

477

478 The accuracy of the normalization method was validated by comparing two distinct methods using the 16S rRNA  
479 and *rpoB* housekeeping genes. The *rpoB* counts were obtained using the translated sequence reads described above  
480 as input for *hmmsearch* (83) against a *rpoB* specific HMM profile (pf04563) from the Pfam-A database (84). Hits  
481 for forward reads (R1) were counted. The correlation between the ARGs normalized to both 16S rRNA, and *rpoB*  
482 was significant ( $R^2 = 0.61$ ,  $p = 4.8 \times 10^{-8}$ ), and we proceeded to use 16S rRNA genes for normalization.

483

484 To study mobile colistin resistance genes (*mcr* genes) and their genetic environments for their theoretical mobility  
485 potential, a subset of the samples (samples BH02, BH48, BFH19, BFH26, BFH41, FH1, and FH9) was assembled  
486 using MEGAHIT (v1.2.8) (85). The minimum contig length was set to 1,000 bp. Contigs harboring mobile colistin  
487 resistance genes were identified using a BLAST (86) search against *mcr* genes in the ResFinder database (n=54).  
488 Similarly, contigs containing class 1 integrons were identified and screened for putative ARGs located in the same  
489 contig.

490

491 Visualizing the genetic environments of contigs harboring *mcr-5.1* colistin ARGs was performed using the *anvi'o*  
492 (v7) (87) platform tool 'anvi-interactive'. For this, single profiles were formed for the samples BH02, BH48,  
493 BFH26, BFH41, and FH9 with the option '-cluster-contigs'. Then, a BLAST nr search of the contigs against the  
494 ResFinder and MobileGeneticElementDatabase was performed. The resulting annotations were imported into the  
495 individual sample profiles containing the contigs database and open reading frames predicted by Prodigal (v2.6.3)  
496 (88) with the command 'anvi-import-functions'. The other predicted genes, sharing the same genetic environment  
497 with the *mcr-5.1* gene, were identified by inspecting the resulting annotations from the BLAST nr search on the  
498 web BLAST 'Graphics' view.

499

500 To simplify the presentation of the relative sum abundance results of the mobile colistin resistance genes, the *mcr*  
501 genes found in ResFinder database (79) were clustered based on 90 % similarity in their sequence identity using  
502 CD-HIT (v4.8.1) (89).

503

504 To visualize the co-localization of multiple ARGs with Bandage (90), the assembled contigs were converted into  
505 assembly graphs using the MEGAHIT command 'megahit\_toolkit contig2fastg'. The Bandage program was run  
506 with a BLAST (86) query of a custom database consisting of ARGs present in the ResFinder database and the  
507 integron sequence-related genes *qacEA* and *intI1*.

508

#### 509 4.4. Statistical analyses

510

511 All statistical analyses described below were performed for all 67 HWW samples described in the 'Sample details'  
512 section and Supporting Information Table S2B. The various wards, clinics, and other units typically had their own  
513 septic tanks or sumps in the Beninese and Burkinabe hospitals. The samples were designated accordingly as  
514 'surgery', 'maternity', 'pediatry', or 'other'. In addition, samples from septic tanks or sumps receiving wastewater  
515 from different sources were designated as 'other'.

516

517 To model the relative sum abundances of ARGs and MGEs in HWWs from Benin, Burkina Faso, and Finland, a  
518 generalized linear model (GLM) with gamma distribution and logarithmic link was fitted to 16S rRNA normalized  
519 count data. The effects of the country, hospital and hospital section on the abundance of ARGs and MGEs were  
520 studied with the function ‘glm’ in the package MASS (v7.3-53.1) (91).

521  
522 The ‘ordinate’ function in the Phyloseq package (v1.30.0) (92) was used for principal coordinate analysis (PCoA)  
523 of resistomes, mobilomes, and taxonomical compositions. Country-related differences in the prevalence of  
524 specific resistance genes were examined using DESeq2 (v1.26.0) (93).

525 The significance of the distinct clustering between the HWWs from different countries drawn using Phyloseq  
526 package (v1.30.0) (92) was tested using pair-wise adonis in the package vegan (v2.5-7) (94) with equal sample  
527 groups. For this, eight samples per country were selected randomly among the HWWs. Unique and shared ARGs  
528 were visualized with the function ‘draw.triple.venn’ in the package VennDiagram (v1.6.20) (95).

529  
530 Pearson correlation coefficients and adjusted *p*-values for each ARG for *intI1* and *qacEA* genes were obtained  
531 with the function ‘corr.test’ in the package psych (v2.1.3) (96) and visualized using package ggcorrplot (v0.1.3)  
532 (97). The correlation tests for different normalization methods (*rpoB* and 16S rRNA) were performed similarly.

533  
534 DESeq2 (v1.26.0) (93) was used to study the differentially abundant ARGs between the HWWs from the three  
535 different countries. ARGs filtered using a variance threshold of 50 were screened for significantly ( $p < 0.05$ )  
536 different relative abundances represented by log-2-fold changes across the HWWs of the three countries in a pair-  
537 wise manner (Benin-Finland, Burkina Faso-Finland). The DESeq2 algorithm is designed for unnormalized count  
538 data from gene expression studies. Thus, to take into account the amount of bacterial DNA in the samples (counts  
539 of detected 16S rRNA), the relative abundances were multiplied by  $10^5$ , and a pseudo count of one was added to  
540 avoid zero values unsuitable for integer transformation. A similar analysis using DESeq2 was performed to study  
541 differentially abundant taxa found in the HWWs from Benin or Burkina Faso versus Finland. The relative  
542 abundances obtained from Metaphlan3 (81) were multiplied with the corresponding 16S rRNA counts for each  
543 sample from the Metaxa2 (82) output. Analyses for differentially abundant genera and species were conducted  
544 separately.

545  
546 Statistical analyses were performed in R (v3.6.3), and the results were visualized using ggplot2 (v3.3.3) (98),  
547 pheatmap (v 1.0.12) (99), and patchwork (v1.1.0.9000) (100). Vector maps were drawn using the package  
548 rnaturalearth (v0.1.0) (101). Detailed descriptions for the bioinformatic and statistical analyses can be viewed on  
549 the project’s GitHub page.

550  
551 Supporting Information

552 The supporting information referred to within the text as Figures S1-S7 and Tables S1-S18 can be found in the  
553 two separate files “Supporting\_Information\_1” (Figures S1-S7, Table S1) and “Supporting\_Information\_2”  
554 (Tables S2-S18).

555  
556 Acknowledgments



557 We thank the staff in all the participating hospitals for their collaboration in sampling. We also thank the laboratory  
558 personnel of the URMAPha at the University of Abomey-Calavi in Benin and the Clinical Research Unit of  
559 Nanoro (CRUN) in Burkina Faso for sample processing. Finally, we acknowledge CSC, IT Center for Science,  
560 Finland, for providing the computational resources for the study.

561

#### 562 Funding Sources

563 This work was supported by the Academy of Finland (Grant Numbers 318643 and 316708) and University of  
564 Helsinki HiLife Grand Challenges funding.

565

#### 566 Ethical permission

567 The ethical permissions for the project were received from “Comité National d’Ethique pour la Recherche en  
568 Santé” under the Health Ministry in Benin and “Comité d’Ethique pour la Recherche en Santé” under the Health  
569 Ministry in Burkina Faso.

570

#### 571 Data availability

572 The data for this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under the  
573 accession number PRJEB47975 and will be public upon article publication in a peer-reviewed journal.

574

#### 575 Code availability

576 All custom codes used for the analyses are available from <https://github.com/melinamarkkanen/AMRIWA> upon  
577 article publication in a peer-reviewed journal.

578

#### 579 References

580

- 581 1. 2015. Global Action Plan On Antimicrobial Resistance WHO Library Cataloguing-in-Publication Data  
582 Global Action Plan on Antimicrobial Resistance <https://doi.org/10.13140/RG.2.2.24460.77448>.
- 583 2. Ventola CL. 2015. The Antibiotic Resistance Crisis Part 1: Causes and Threats. *Physical Therapy*  
584 40:277–83.
- 585 3. Laxminarayan R, Duse A, Wattal C, Zaidi AKM, Wertheim HFL, Sumpradit N, Vlieghe E, Hara GL,  
586 Gould IM, Goossens H, Greko C, So AD, Bigdeli M, Tomson G, Woodhouse W, Ombaka E, Peralta  
587 AQ, Qamar FN, Mir F, Kariuki S, Bhutta ZA, Coates A, Bergstrom R, Wright GD, Brown ED, Cars O.  
588 2013. Antibiotic resistance—the need for global solutions. *Lancet Infect Dis* 13:1057–1098.
- 589 4. Jacobs J, Hardy L, Semret M, Lunguya O, Phe T, Affolabi D, Yansouni C, Vandenberg O. 2019.  
590 Diagnostic Bacteriology in District Hospitals in Sub-Saharan Africa: At the Forefront of the  
591 Containment of Antimicrobial Resistance. *Frontiers in Medicine* 6:205.
- 592 5. Belachew SA, Hall L, Selvey LA. 2021. Non-prescription dispensing of antibiotic agents among  
593 community drug retail outlets in Sub-Saharan African countries: a systematic review and meta-analysis.  
594 *Antimicrobial Resistance and Infection Control*. BioMed Central Ltd.
- 595 6. Butcher A, Cañada JA, Sariola S. 2021. How to make noncoherent problems more productive: Towards  
596 an AMR management plan for low resource livestock sectors. *Humanities and Social Sciences*  
597 *Communications* 8:287.
- 598 7. Van TTH, Yidana Z, Smooker PM, Coloe PJ. 2020. Antibiotic use in food animals worldwide, with a  
599 focus on Africa: Pluses and minuses. *Journal of Global Antimicrobial Resistance* 20:170–177.
- 600 8. Hu Y, Gao GF, Zhu B. 2017. The antibiotic resistome: gene flow in environments, animals and human  
601 beings. *Frontiers of Medicine* 11:161–168.
- 602 9. Tacconelli E, Carrara E, Savoldi A, Kattula D, Burkert F. 2017. Global Priority List Of Antibiotic-  
603 Resistant Bacteria To Guide Research, Discovery, And Development Of New Antibiotics.
- 604 10. Martínez JL, Coque TM, Baquero F. 2014. What is a resistance gene? Ranking risk in resistomes.  
605 *Nature Reviews Microbiology* 13:116–123.

- 606 11. Zhang A-N, Gaston JM, Dai CL, Zhao S, Poyet M, Groussin M, Yin X, Li L-G, van Loosdrecht MCM,  
607 Topp E, Gillings MR, Hanage WP, Tiedje JM, Moniz K, Alm EJ, Zhang T. An omics-based framework  
608 for assessing the health risk of antimicrobial resistance genes [https://doi.org/10.1038/s41467-021-](https://doi.org/10.1038/s41467-021-25096-3)  
609 25096-3.
- 610 12. Karkman A, Berglund F, Flach CF, Kristiansson E, Larsson DGJ. 2020. Predicting clinical resistance  
611 prevalence using sewage metagenomic data. *Communications Biology* 3:711.
- 612 13. Gillings MR, Gaze WH, Pruden A, Smalla K, Tiedje JM, Zhu Y-G. 2015. Using the class 1 integron-  
613 integrase gene as a proxy for anthropogenic pollution. *The ISME Journal* 9:1269–1279.
- 614 14. Li Y, Yang X, Zhang J, Yang S, Zhang S, Chen M, Xue L, Ding Y, Zeng H, Gu Q, Zhang Y, Wei X,  
615 Wang J, Wu Q. 2021. Molecular characterisation of antimicrobial resistance determinants and class 1  
616 integrons of *Salmonella enterica* subsp. *enterica* serotype Enteritidis strains from retail food in China.  
617 *Food Control* 128:108191.
- 618 15. Gillings MR. 2014. Integrons: Past, Present, and Future. *Microbiology and Molecular Biology Reviews*  
619 78:257–77.
- 620 16. Gillings M, Boucher Y, Labbate M, Holmes A, Krishnan S, Holley M, Stokes HW. 2008. The Evolution  
621 of Class 1 Integrons and the Rise of Antibiotic Resistance. *JOURNAL OF BACTERIOLOGY*  
622 190:5095–5100.
- 623 17. Chen B, Han J, Dai H, Jia P. 2021. Biocide-tolerance and antibiotic-resistance in community  
624 environments and risk of direct transfers to humans: Unintended consequences of community-wide  
625 surface disinfecting during COVID-19? *Environmental Pollution* 283:117074.
- 626 18. de Angelis G, Fiori B, Menchinelli G, Tiziana D'Inzeo &, Flora &, Liotti M, Morandotti GA,  
627 Sanguinetti M, Posteraro B, Spanu T. 2007. Incidence and antimicrobial resistance trends in  
628 bloodstream infections caused by ESKAPE and *Escherichia coli* at a large teaching hospital in Rome, a  
629 9-year analysis (2007-2015). *European Journal of Clinical Microbiology & Infectious Diseases*  
630 37:1627–1636.
- 631 19. Alemayehu T, Hailemariam M. 2020. Prevalence of vancomycin-resistant enterococcus in Africa in one  
632 health approach: a systematic review and meta-analysis. *Scientific Reports* 10:20542.
- 633 20. Karkman A, Do TT, Walsh F, Virta MPJ. 2018. Antibiotic-Resistance Genes in Waste Water. *Trends in*  
634 *Microbiology* 26:220–228.
- 635 21. World Health Organization. 2021. Global Antimicrobial Resistance and Use Surveillance System  
636 (GLASS) Report 2021. Geneva.
- 637 22. World Health Organization. 2018. WHO report on surveillance of antibiotic consumption: 2016-2018  
638 early implementation.
- 639 23. Suay-García, Pérez-Gracia. 2019. Present and Future of Carbapenem-resistant Enterobacteriaceae  
640 (CRE) Infections. *Antibiotics* 8:122.
- 641 24. Poirel L, Marqué S, Héritier C, Segonds C, Chabanon G, Nordmann P. 2005. OXA-58, a novel class D  
642  $\beta$ -lactamase involved in resistance to carbapenems in *Acinetobacter baumannii*. *Antimicrobial Agents*  
643 *and Chemotherapy* 49:202–208.
- 644 25. Poirel L, Weldhagen GF, Naas T, de Champs C, Dove MG, Nordmann P. 2001. GES-2, a class A  $\beta$ -  
645 lactamase from *Pseudomonas aeruginosa* with increased hydrolysis of imipenem. *Antimicrobial Agents*  
646 *and Chemotherapy* 45:2598–2603.
- 647 26. Nordmann P, Poirel L. 2013. The difficult-to-control spread of carbapenemase producers among  
648 Enterobacteriaceae worldwide. *Clinical Microbiology and Infection* 20:821–830.
- 649 27. Löfmark S, Edlund C, Erik Nord C. 2010. Metronidazole Is Still the Drug of Choice for Treatment of  
650 Anaerobic Infections. *Clinical Infectious Diseases* 1:S16–S23.
- 651 28. Alauzet C, Lozniewski A, el Ene Marchandin H. 2018. Metronidazole resistance and nim genes in  
652 anaerobes: A review. *Anaerobe* 55:40–53.
- 653 29. Falagas ME, Kasiakou Sofia. K. 2005. Colistin: The Revival of Polymyxins for the Management of  
654 Multidrug-Resistant Gram-Negative Bacterial Infections. *Clinical Infectious Diseases* 40:1333–41.
- 655 30. Ling Z, Yin W, Shen Z, Wang Y, Shen J, Walsh TR. 2020. Epidemiology of mobile colistin resistance  
656 genes mcr-1 to mcr-9. *Journal of Antimicrobial Chemotherapy* 75:3087–3095.
- 657 31. Anyanwu MU, Odilichukwu C, Okpala R, Chah KF, Shoyinka VS. 2021. Prevalence and Traits of  
658 Mobile Colistin Resistance Gene Harbouring Isolates from Different Ecosystems in Africa. *BioMed*  
659 *Research International* 22:6630379.
- 660 32. Collignon P, Beggs JJ, Walsh TR, Gandra S, Laxminarayan R. 2018. Anthropological and  
661 socioeconomic factors contributing to global antimicrobial resistance: a univariate and multivariable  
662 analysis. *Lancet Planet Health* 2:e398-405.
- 663 33. van Boeckel TP, Pires J, Silvester R, Zhao C, Song J, Criscuolo NG, Gilbert M, Bonhoeffer S,  
664 Laxminarayan R. 2019. Global trends in antimicrobial resistance in animals in low-and middle-income  
665 countries. *Science* 365:eaaw1944.

- 666 34. Iskandar K, Molinier L, Hallit S, Sartelli M, Craig Hardcastle T, Haque M, Lugova H, Dhingra S,  
667 Sharma P, Islam S, Mohammed I, Naina Mohamed I, Abi Hanna P, el Hajj S, Adilla Hayat Jamaluddin  
668 N, Salameh P, Roques C. 2020. Surveillance of antimicrobial resistance in low- and middle-income  
669 countries: a scattered picture. *Antimicrob Resist Infect Control* 10:63.
- 670 35. Ambler RP. 1980. The structure of beta-lactamases. *Philosophical transactions of the Royal Society of*  
671 *London Series B, Biological sciences* 289:321–331.
- 672 36. Bernabé KJ, Langendorf C, Ford N, Ronat J-B, Murphy RA. 2017. Antimicrobial resistance in West  
673 Africa: a systematic review and meta-analysis. *International Journal of Antimicrobial Agents* 50:629–  
674 639.
- 675 37. Mahamat OO, Kempf M, Lounnas M, Tidjani A, Hide M, Benavides JA, Carrière C, Bañuls A-L, Jean-  
676 Pierre H, Ouedraogo A-S, Dumont Y, Godreuil S. 2021. Epidemiology and prevalence of extended-  
677 spectrum  $\beta$ -lactamase-and carbapenemase-producing Enterobacteriaceae in humans, animals and the  
678 environment in West and Central Africa. *International Journal of Antimicrobial Agents* 57:106203.
- 679 38. Hendriksen RS, Munk P, Njage P, van Bunnik B, McNally L, Lukjancenko O, Röder T, Nieuwenhuijse  
680 D, Pedersen SK, Kjeldgaard J, Kaas RS, Clausen PTL, Vogt JK, Leekitcharoenphon P, van de Schans  
681 MGM, Zuidema T, de Roda Husman AM, Rasmussen S, Petersen B, Amid C, Cochrane G, Sicheritz-  
682 Ponten T, Schmitt H, Alvarez JRM, Aidara-Kane A, Pamp SJ, Lund O, Hald T, Woolhouse M,  
683 Koopmans MP, Vigre H, Petersen TN, Aarestrup FM. 2019. Global monitoring of antimicrobial  
684 resistance based on metagenomics analyses of urban sewage. *Nature Communications* 10:1124.
- 685 39. Pärnänen KMM, Narciso-da-Rocha C, Kneis D, Berendonk TU, Cacace D, Do TT, Elpers C, Fatta-  
686 Kassinos D, Henriques I, Jaeger T, Karkman A, Martinez JL, Michael SG, Michael-Kordatou I,  
687 O'Sullivan K, Rodriguez-Mozaz S, Schwartz T, Sheng H, Sørum H, Stedtfeld RD, Tiedje JM, Giustina  
688 SV della, Walsh F, Vaz-Moreira I, Virta M, Manaia CM. 2019. Antibiotic resistance in European  
689 wastewater treatment plants mirrors the pattern of clinical antibiotic resistance prevalence. *Science*  
690 *Advances* 5:eaau9124.
- 691 40. European Centre for Disease Prevention and Control. 2020. Antimicrobial resistance in the EU/EEA  
692 (EARS-Net) - Annual Epidemiological Report 2019.
- 693 41. Buelow E, Rico A, Gaschet M, Lourenço J, Kennedy SP, Wiest L, Ploy M-C, Dagot C. 2020. Hospital  
694 discharges in urban sanitation systems: Long-term monitoring of wastewater resistome and microbiota  
695 in relationship to their eco-exposome. *Water Research X* 7:100045.
- 696 42. Thomas AM, Segata N. Multiple levels of the unknown in microbiome research  
697 <https://doi.org/10.1186/s12915-019-0667-z>.
- 698 43. Naas T, Oueslati S, Bonnin RA, Laura Dabos M, Zavala A, Dortet L, Retailleau P, Iorga BI, emy  
699 Bonnin RA. 2017. Beta-lactamase database (BLDB)-structure and function. *Journal of Enzyme*  
700 *Inhibition and Medicinal Chemistry* 32:917–919.
- 701 44. Thompson L, Sanders JG, McDonald D, Amir A, Ladau J, Locey KJ, Prill robert J, Tripathi A, Gibbons  
702 S, Ackermann G, Navas-molina JA, Janssen S, Kopylova E, Vázquez-baeza Y, González A, Morton JT,  
703 Mirarab S, Zech Xu Z, Jiang L, Haroon mohamed F, Kanbar J, Zhu Q, Jin Song S, Kosciolk T,  
704 Bokulich nicholas A, Lefler J, Brislawn colin J, Humphrey G. 2017. A communal catalogue reveals  
705 Earth's multiscale microbial diversity. *Nature* 551:457–463.
- 706 45. Oyem IM, Oyem HH, Atuanya. 2020. Molecular characterisation of bacteria strains of septic tank  
707 sewage samples from related sites in Delta and Edo States of Nigeria using 16S rRNA denaturing  
708 gradient gel electrophoresis (DGGE) 14:132–138.
- 709 46. Pärnänen K, Karkman A, Hultman J, Lyra C, Bengtsson-Palme J, Larsson DGJ, Rautava S, Isolauri E,  
710 Salminen S, Kumar H, Satokari R, Virta M. 2018. Maternal gut and breast milk microbiota affect infant  
711 gut antibiotic resistome and mobile genetic elements. *Nature Communications* 9:389.
- 712 47. Laursen MF, Laursen RP, Larnkjaer A, Mølgaard C, Michaelsen KF, Frøkiaer H, Bahl MI, Licht TR.  
713 2017. Faecalibacterium Gut Colonization Is Accelerated by Presence of Older Siblings. *mSphere*  
714 2:e00448-17.
- 715 48. Collis CM, Hall RM. 1995. Expression of antibiotic resistance genes in the integrated cassettes of  
716 integrons. *Antimicrobial Agents and Chemotherapy* 39:155–162.
- 717 49. Périchon B, Goussard S, Walewski V, Krizova L, Cerqueira G, Murphy C, Feldgarden M, Wortman J,  
718 Clermont D, Nemeč A, Courvalin P. 2014. Identification of 50 class D  $\beta$ -lactamases and 65  
719 *Acinetobacter*-derived cephalosporinases in *Acinetobacter* spp. *Antimicrobial Agents and*  
720 *Chemotherapy* 58:936–949.
- 721 50. Figueiredo S, Bonnin RA, Poirel L, Duranteau J, Nordmann P. 2012. Identification of the naturally  
722 occurring genes encoding carbapenem-hydrolysing oxacillinases from *Acinetobacter haemolyticus*,  
723 *Acinetobacter johnsonii*, and *Acinetobacter calcoaceticus*. *Clin Microbiol Infect* 18:907–913.
- 724 51. Massidda O, Rossolini GM, Satta G. 1991. The *Aeromonas hydrophila* cphA gene: Molecular  
725 heterogeneity among class B metallo- $\beta$ -lactamases. *Journal of Bacteriology* 173:4611–4617.

- 726 52. Cox G, Wright GD. 2013. Intrinsic antibiotic resistance: Mechanisms, origins, challenges and solutions.  
727 *International Journal of Medical Microbiology* 303:287–292.
- 728 53. Araújo S, Sousa M, Tacão M, Baraúna RA, Silva A, Ramos R, Alves A, Manaia CM, Henriques I.  
729 2021. Carbapenem-resistant bacteria over a wastewater treatment process: Carbapenem-resistant  
730 Enterobacteriaceae in untreated wastewater and intrinsically-resistant bacteria in final effluent. *Science*  
731 *of the Total Environment* 782:146892.
- 732 54. Zieliński W, Korzeniewska E, Harnisz M, Drzymala J, Felis E, Bajkacz S. 2021. Wastewater treatment  
733 plants as a reservoir of integrase and antibiotic resistance genes – An epidemiological threat to workers  
734 and environment. *Environment International* 156:106641.
- 735 55. Pal C, Bengtsson-Palme J, Kristiansson E, Larsson DGJ. 2015. Co-occurrence of resistance genes to  
736 antibiotics, biocides and metals reveals novel insights into their co-selection potential. *BMC Genomics*  
737 16.
- 738 56. Gillings MR, Xuejun D, Hardwick SA, Holley MP, Stokes HW. 2009. Gene cassettes encoding  
739 resistance to quaternary ammonium compounds: a role in the origin of clinical class 1 integrons? *The*  
740 *ISME Journal* 3:209–215.
- 741 57. Olalekan A, Onwugamba F, Iwalokun B, Mellmann A, Becker K, Schaumburg F. 2020. High proportion  
742 of carbapenemase-producing *Escherichia coli* and *Klebsiella pneumoniae* among extended-spectrum  $\beta$ -  
743 lactamase-producers in Nigerian hospitals. *Journal of Global Antimicrobial Resistance* 21:8–12.
- 744 58. Manenzhe RI, Zar HJ, Nicol MP, Kaba M. 2015. The spread of carbapenemase-producing bacteria in  
745 Africa: a systematic review. *Journal of Antimicrobial Chemotherapy* 70:23–40.
- 746 59. van Duin D, Doi Y. 2017. The global epidemiology of carbapenemase-producing Enterobacteriaceae.  
747 *Virulence* 8:460–469.
- 748 60. Mushi MF, Mshana SE, Imirzalioglu C, Bwanga F. 2014. Carbapenemase Genes among Multidrug  
749 Resistant Gram Negative Clinical Isolates from a Tertiary Hospital in Mwanza, Tanzania. *BioMed*  
750 *Research International* 2014:303104.
- 751 61. Yehouenou C, Bogaerts B, Vanneste K, Roosens NHC, de Keersmaecker SCJ, Marchal K, Affolabi D,  
752 Soleimani R, Rodriguez-Villalobos H, van Bambeke F, Dalleur O, Simon A. 2021. First detection of a  
753 plasmid-encoded New-Delhi metallo-beta-lactamase-1 (NDM-1) producing *Acinetobacter baumannii*  
754 using whole genome sequencing, isolated in a clinical setting in Benin. *Annals of Clinical Microbiology*  
755 *and Antimicrobials* 20:5.
- 756 62. Mutai WC, Mureithi MW, Anzala O, Revathi G, Kullin B, Burugu M, Kyany'a C, Odoyo E, Otieno P,  
757 Musila L. 2021. High Prevalence of Multidrug-Resistant *Clostridioides difficile* Following Extensive  
758 Use of Antimicrobials in Hospitalized Patients in Kenya. *Frontiers in Cellular and Infection*  
759 *Microbiology* 10:897.
- 760 63. Olumuyiwa Olaitan A, Dandachi I, Alexandra Baron S, Daoud Z, Morand S, Rolain J-M. 2021. Banning  
761 colistin in feed additives: a small step in the right direction. *The Lancet Infectious Diseases* 21:29–30.
- 762 64. Olowo-Okere A, Yacouba A. 2020. Review Molecular mechanisms of colistin resistance in Africa: A  
763 systematic review of literature. *Germs* 10:367–379.
- 764 65. Kneis D, Berendonk TU, Heß S. 2019. High prevalence of colistin resistance genes in German  
765 municipal wastewater. *Science of the Total Environment* 694:133454.
- 766 66. Ngbede EO, Poudel A, Kalalah A, Yang Y, Adekanmbi F, Adikwu AA, Adamu AM, Mamfe LM,  
767 Daniel ST, Useh NM, Kwaga JKP, Adah MI, Kelly P, Butaye P, Wang C. 2020. Identification of mobile  
768 colistin resistance genes (*mcr-1.1*, *mcr-5* and *mcr-8.1*) in Enterobacteriaceae and *Alcaligenes faecalis* of  
769 human and animal origin, Nigeria. *International Journal of Antimicrobial Agents* 56:106108.
- 770 67. Borowiak M, Fischer J, Hammerl JA, Hendriksen RS, Szabo I, Malorny B. 2017. Identification of a  
771 novel transposon-associated phosphoethanolamine transferase gene, *mcr-5*, conferring colistin  
772 resistance in d-tartrate fermenting *Salmonella enterica* subsp. *enterica* serovar Paratyphi B. *Journal of*  
773 *Antimicrobial Chemotherapy* 72:3317–3324.
- 774 68. Borowiak M, Hammerl JA, Deneke C, Fischer J, Szabo I, Malorny B. 2019. Characterization of *mcr-5*-  
775 Harboring *Salmonella enterica* subsp. *enterica* Serovar Typhimurium Isolates from Animal and Food  
776 Origin in Germany. *Antimicrobial Agents and Chemotherapy* 63:e00063-19.
- 777 69. Guo S, Tay MYF, Thu AK, Seow KLG, Zhong Y, Ng LC, Schlundt J. 2019. Conjugative IncX1 plasmid  
778 harboring colistin resistance gene *mcr-5.1* in *Escherichia coli* isolated from chicken rice retailed in  
779 Singapore. *Antimicrobial Agents and Chemotherapy* 63:e01043-19.
- 780 70. Peace Hounkpe S, Codjo Adjovi E, Crapper M, Awuah E. 2014. Wastewater Management in Third  
781 World Cities: Case Study of Cotonou. *Benin Journal of Environmental Protection* 5:387–399.
- 782 71. Bougnom BP, Zongo C, McNally A, Ricci V, Etoa FX, Thiele-Bruhn S, Piddock LJV. 2019.  
783 Wastewater used for urban agriculture in West Africa as a reservoir for antibacterial resistance  
784 dissemination. *Environmental Research* 168:14–24.

- 785 72. Bougnom BP, Thiele-Bruhn S, Ricci V, Zongo C, Piddock LJV. 2020. Raw wastewater irrigation for  
786 urban agriculture in three African cities increases the abundance of transferable antibiotic resistance  
787 genes in soil, including those encoding extended spectrum  $\beta$ -lactamases (ESBLs). *Science of the Total*  
788 *Environment* 698:134201.
- 789 73. Totin HS, Ernest A, Odoulami L, Edoth PA, Boukari M, Boko M. 2013. Groundwater pollution and the  
790 safe water supply challenge in Cotonou town, Benin (West Africa). *IAHS-AISH Proceedings and*  
791 *Reports* 361:191–196.
- 792 74. Mölder F, Jablonski KP, Letcher B, Hall MB, Tomkins-Tinch CH, Sochat V, Forster J, Lee S,  
793 Twardziok SO, Kanitz A, Wilm A, Holtgrewe M, Rahmann S, Nahnsen S, Köster J. 2021. Sustainable  
794 data analysis with Snakemake. *F1000Research* 10:33.
- 795 75. Andrews S. 2010. FastQC A quality control tool for high throughput sequence data  
796 <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
- 797 76. Ewels P, Ns Magnusson M, Lundin S, Aller MK. Data and text mining MultiQC: summarize analysis  
798 results for multiple tools and samples in a single report. *Bioinformatics* 32:3047–3048.
- 799 77. Marcel M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads.  
800 *EMBnet.journal* 17:10–12.
- 801 78. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9:357–  
802 359.
- 803 79. Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Cattoir V, Philippon A, Allesoe RL, Rebelo  
804 AR, Florensa AF, Fagelhauer L, Chakraborty T, Neumann B, Werner G, Bender JK, Stingl K, Nguyen  
805 M, Coppens J, Xavier BB, Malhotra-Kumar S, Westh H, Pinholt M, Anjum MF, Duggett NA, Kempf I,  
806 Nykäsenoja S, Olkkola S, Wiczorek K, Amaro A, Clemente L, Mossong J, Losch S, Ragimbeau C,  
807 Lund O, Aarestrup FM. 2020. ResFinder 4.0 for predictions of phenotypes from genotypes. *Journal of*  
808 *Antimicrobial Chemotherapy* 75:3491–3500.
- 809 80. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009.  
810 The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079.
- 811 81. Beghini F, Mciver LJ, Blanco-mi A, Dubois L, Asnicar F, Maharjan S, Mailyan A, Manghi P, Scholz  
812 M, Thomas AM, Valles-colomer M, Weingart G, Zhang Y, Zolfo M, Huttenhower C, Franzosa EA,  
813 Segata N. 2021. Integrating taxonomic, functional, and strain-level profiling of diverse microbial  
814 communities with bioBakery 3. *Elife* 4:1–42.
- 815 82. Bengtsson-Palme J, Hartmann M, Eriksson KM, Pal C, Thorell K, Larsson DGJ, Nilsson RH. 2015.  
816 metaxa2: Improved identification and taxonomic classification of small and large subunit rRNA in  
817 metagenomic data. *Molecular Ecology Resources* 15:1403–1414.
- 818 83. Eddy SR. 2011. Accelerated Profile HMM Searches. Citation: Eddy SR 7:1002195.
- 819 84. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar  
820 GA, Smart A, Sonnhammer ELL, Hirsh L, Paladin L, Piovesan D, Tosatto SCE, Finn RD. 2018. The  
821 Pfam protein families database in 2019. *Nucleic Acids Research* 47:427–432.
- 822 85. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. 2015. MEGAHIT: an ultra-fast single-node solution for  
823 large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31:1674–1676.
- 824 86. Altschup SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic Local Alignment Search Tool. *J*  
825 *Mol Biol* 215:403–410.
- 826 87. Murat Eren A, Kiefl E, Shaiber A, Veseli I, Miller SE, Schechter MS, Fink I, Pan JN, Yousef M,  
827 Fogarty EC, Trigodet F, Watson AR, Esen ÖC, Moore RM, Clayssen Q, Lee MD, Kivenson V, Graham  
828 ED, Merrill BD, Karkman A, Blankenberg D, Eppley JM, Sjödin A, Scott JJ, Vázquez-Campos X,  
829 McKay LJ, McDaniel EA, R Stevens SL, Anderson RE, Fuessel J, Fernandez-Guerra A, Maignien L,  
830 Delmont TO, Willis AD. 2021. Community-led, integrated, reproducible multi-omics with anvi'o.  
831 *Nature Microbiology* 6:3–6.
- 832 88. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene  
833 recognition and translation initiation site identification. *BMC bioinformatics* 11.
- 834 89. Fu L, Niu B, Zhu Z, Wu S, Li W. 2012. Sequence analysis CD-HIT: accelerated for clustering the next-  
835 generation sequencing data 28:3150–3152.
- 836 90. Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of de novo genome  
837 assemblies. *Bioinformatics* 31:3350–3352.
- 838 91. Venables WN, Ripley B. 2002. *Modern Applied Statistics with S*.
- 839 92. McMurdie PJ, Holmes S. 2013. Phyloseq: An R Package for Reproducible Interactive Analysis and  
840 Graphics of Microbiome Census Data. *PLoS ONE* 8.
- 841 93. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq  
842 data with DESeq2. *Genome Biology* 15:550.

- 843 94. Oksanen J, Blanchet G, Friendly M, Kindt R, Legendre P, McGlenn D, Minchin PR, O'Hara RB,  
844 Simpson GL, Solymos P, Stevens MHM, Szocs E, Wagner H. 2020. *vegan*: Community Ecology  
845 Package.
- 846 95. Chen H. 2018. *VennDiagram*: Generate High-Resolution Venn and Euler Plots. R package version  
847 1.6.20.
- 848 96. Revelle W. 2021. *psych*: Procedures for Psychological, Psychometric, and Personality Research.
- 849 97. Kassambara A. 2019. *ggcorrplot*: Visualization of a Correlation Matrix using "ggplot2."
- 850 98. Wickham H. 2016. *ggplot2*: Elegant graphics for data analysis. Springer-Verlag New York.
- 851 99. Kolde R. 2019. *pheatmap*: Pretty Heatmaps. R package version 1.0.12.
- 852 100. Pedersen TL. 2020. *patchwork*: The Composer of Plots. R package version 1.1.1.
- 853 101. South A. *rnaturalearth*: World Map Data from Natural Earth. R package version 0.1.0.
- 854