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

Melina A. Markkanen[1]*, Kaisa Haukka[1,6], Katariina M. M. Pärnänen[1], Victorien Tamegnon Dougnon[2], Isidore Juste O. Bonkoungou[3], Zakaria Garba[4], Halidou Tinto[4], Anniina Sarekoski[1], Antti Karkman[1], Anu Kantele[5,6]†, Marko P. J. Virta[1]*†

[1]Department of Microbiology, University of Helsinki, Finland
[2]Research Unit in Applied Microbiology and Pharmacology of Natural Substances, Polytechnic School of Abomey-Calavi, University of Abomey-Calavi, Benin
[3]Department of Biochemistry and Microbiology, University Joseph KI-ZERBO, Ouagadougou, Burkina Faso
[4]Clinical Research Unit Nanoro, Institute for Research in Health Sciences, National Center for Scientific and Technological Research, Burkina Faso
[5]Meilahti Infectious Diseases and Vaccine Research Center, MeiVac, Department of Infectious Diseases, University of Helsinki and Helsinki University Hospital, Helsinki, Finland
[6]Human Microbiome Research Program, Faculty of Medicine, University of Helsinki, Helsinki, Finland

Author information

Corresponding authors:
Marko P. J. Virta, marko.virta@helsinki.fi
Melina A. Markkanen, melina.markkanen@helsinki.fi

† Equal contribution

ABSTRACT

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

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

Keywords

Antimicrobial resistance (AMR)
West Africa
hospital wastewater (HWW)
carbapenemase
colistin resistance
metagenomes

Synopsis

Metagenomic analysis of HWWs revealed high abundances of ARGs and their potential transmission to other environments in Benin and Burkina Faso.

1. Introduction

The AMR crisis is a global issue with multifaceted effects on human, animal, and environmental health, and it comes with substantial economic losses[1]. Resistance to almost all antibiotics, including the last resort regimen reserved to treat the most severe infections, has tended to emerge only a couple of years after introducing a new antibiotic to the global market[2,3]. Due to limitations in diagnostic testing in low-resource settings in low-income countries (LICs), broad-spectrum antibiotics are often prescribed empirically without microbiological verification of the causative pathogen or its sensitivity to different antibiotics[4].

Bacterial resistance to antimicrobial compounds may be intrinsic, or acquired from other bacteria or the surrounding environment via horizontal gene transfer of mobile genetic elements (MGE)[5–7]. While both intrinsic and acquired resistance contribute to the total resistance burden, acquired resistance is considered to have a greater impact on the current AMR problem induced by human activity[7]. Class 1 integrons are strongly linked to the dissemination of clinically relevant acquired ARGs[8,9]. These integrons are known for their ability to 'collect' several genes by site-specific recombination under environmental stress, such as the presence of antimicrobial residues[10]. The resulting integron-carried gene cassettes, consisting of several ARGs, are expressed by a single integron-associated promoter[11]. Class 1 integron gene cassettes can become transmitted to new hosts, for example, via plasmids[12]. The gene encoding the integron integrase (*intI*) among other associated genes, is involved in the recombination of incoming genetic material into the gene cassette, and *intI1* is typically used as a marker for class 1 integrons[9,10]. Over time, quaternary ammonium compound (QAC) resistance genes, such as *qacEΔ,* have established their position in the structure of a typical class 1 integron gene cassette whose spread is induced by human activity[10,13]. It has been speculated that increasing the use of such disinfectant compounds in hospitals and community environments, especially during the COVID-19 pandemic, would further exacerbate the problem of multidrug resistance mediated by class 1 integrons[14].

The high-risk ESKAPEEc bacteria[6] (*Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa, Enterobacter* spp. and *Escherichia coli*)[15] are top priority pathogens for AMR research and the development of new antibiotics[6,16]. They can feature an impressive resistance profile in terms of both intrinsic and acquired ARGs and often exhibit multidrug resistant phenotypes. For instance, hospital-associated strains of methicillin-resistant *S. aureus* (MRSA) carrying the *mec*[17] gene that encodes a beta-lactamase and vancomycin-resistant *Enterococci* (VRE) carrying the *van* gene[16] are clinically important pathogens in hospitals and are both associated with increased mortality[1,18]. However, investigating only bacterial isolates may result in limited understanding of ARGs occurring in a particular hospital. In contrast, shotgun metagenomic sequencing enables the analysis of resistomes and mobilomes (all ARGs or MGEs, respectively) as well as the microbial communities present in the studied environment[19]. Hence, as opposed to PCR based methods, there is no requirement for a priori knowledge of the resistome[19]. Metagenomic approaches have been applied to provide a more comprehensive picture of the AMR situation in wastewaters[19,20] as an alternative to current AMR surveillance practices focusing on clinical isolates[21].

As a consequence of the increased prevalence of extended-spectrum beta-lactamase-producing *Enterobacteriaceae* (ESBL-PE), the use of broad-spectrum antibiotics has increased in clinical practise[2,21,22]. Tracking down the spread of ESBL genes is, in many situations, no longer feasible or relevant, and attention has been directed to plasmid-mediated genes that encode carbapenemases. The carbapenemases *bla*GES, *bla*IMP, *bla*KPC, *bla*NDM, *bla*OXA-48, *bla*OXA-58, and *bla*VIM carried by plasmids have emerged and spread around the world during the past three decades[23–26]. However, only a subset of all the currently known *bla*GES genes confer resistance to carbapenems, namely, *bla*GES-2, -4, -5, -6, -14, -16, and -20[26]. *bla*GES-5 is one of the fastest and most widely dispersed *bla*GES genes among *Enterobacteriaceae*[26].

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

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

As hospital wastewater (HWW) from healthcare facilities is considered to be at the frontline of AMR emergence and spread[41], we set out to investigate the AMR situation and the characteristics of the resistomes in nine hospitals

122 in two West African countries, Benin and Burkina Faso, where prior data was scarce[21]. For comparison, we
123 analyzed samples from four hospitals in Finland, where the level of AMR was expected to be low[40]. In this study,
124 we applied shotgun metagenomic sequencing as it allows the analysis of the total catalog of genes present in the
125 studied environment.
126
127 *2. Material and Methods*
128
129 *2.1. Sample description*
130
131 Hospital wastewater (HWW) samples were collected in Benin from four (hospitals A-D) and in Burkina Faso
132 from five different hospitals (hospitals F-J) in November and December of 2019. In Finland, HWW samples were
133 collected from four different hospitals (hospitals K-N) in January of 2020. In Burkina Faso, the samples were
134 mainly from septic tanks or sewers in the hospital area. In Benin, none of the hospitals was connected to a sewer
135 system, and the samples were from septic tanks or sumps (unstructured wastewater wells), which were never
136 emptied to our knowledge. In most cases the toilet water was not directed into these sumps. For comparison, some
137 other types of samples were included: samples from river waters used for drinking in a remote countryside village
138 in the community of Savalou in central Benin (n=3); river water receiving treated wastewater in Burkina Faso
139 (n=2); and fecal samples of in-patients in Benin (n=4). In addition, the following samples were analyzed; well
140 water, treated tap water, water for handwashing, water in a municipality channel receiving biologically treated
141 wastewater, water and sediment from a street gutter next to the hospital, an empty HWW septic tank, and the soil
142 between HWW septic tanks. The four fecal samples were obtained from patients staying at hospital B with
143 diagnoses mainly associated with musculoskeletal injuries. No clinical data concerning bacterial infections were
144 reported for these patients, and we have no information on their possible use of antibiotics. Patient age varied
145 from 24 to 50 years old, and at the time of sampling, patients had stayed in the hospital for one to two months.
146 All sample details are provided in the Supplementary Material, Table S1A. Selected pictures and a map indicating
147 sample collection regions are shown in Figures S1 and S2 in the Supplementary Material. Letters A–N were used
148 to denote the hospitals/sampling locations.
149
150 *2.2. DNA extraction and metagenomic sequencing*
151
152 Water samples were collected into 1-liter bottles, which were transported to a laboratory on ice. They were kept
153 at +4 °C until processed within 24 hours. A volume of 50–100 ml was filtered through a 0.2 µm polycarbonate
154 filter (Whatman™, GE Healthcare Life Sciences) using a portable vacuum pump (Millivac-Mini Vacuum Pump
155 XF54, Millipore, Merck). DNA was extracted from the filters using the Qiagen DNeasy PowerWater DNA Kit
156 for the water samples and Qiagen DNeasy PowerSoil DNA Kit for the soil samples, following manufacturer
157 instructions. The concentration and quality of the extracted DNA were determined with a NanoDrop
158 spectrophotometer. Altogether 96 samples were subjected to shotgun metagenomic sequencing using Illumina
159 Novaseq5000 with Nextera XT library preparation at the Institute of Biotechnology, University of Helsinki. The
160 sequencing run failed for three samples, one from each country (Supplementary Material, Table S1, with the
161 samples shown in red).

2.3. Metagenomic analysis

For the analyses comparing the HWWs from different countries, a subset of 67 samples was created: Benin, n = 25; Burkina Faso, n = 34; Finland, n = 8. This subset of samples is shown in the Supplementary Material, Table S1B. The various wards, clinics, and other units typically had their own septic tanks or sumps in the Beninese and Burkinabe hospitals. The samples were designated accordingly as 'surgery', 'maternity', 'pediatry', or 'other'. In addition, samples from septic tanks or sumps receiving wastewater from different sources were designated as 'other'. The Finnish samples were from the main hospital drains collecting water from various hospital sections. In addition, some technical and biological replicates (e.g., multiple samples collected from the same septic tank, sump, or drain) were collected and analyzed. These were excluded from the statistical analyses when comparing HWWs from different countries.

All samples were analyzed using an in-house Snakemake workflow. Snakemake enables reproducible analysis for large datasets through stable version control of the bioinformatic programs used[42]. The program versions and parameters applied for each program can be viewed on the project's GitHub page.

Briefly, the quality control steps included in the workflow were performed using FastQC[43] and MultiQC[44], with adapter and low-quality read removal using Cutadapt[45]. Nucleotide sequence reads were mapped using Bowtie2[46] against the ResFinder database (version 3.2, downloaded on 2020-06-28)[47]. The reads were sorted and filtered using SAMtools[48] such that the reads mapping as pair or alone were calculated as a single count. Mobile genetic elements were identified by mapping the reads against the MobileGeneticElementDatabase (https://github.com/KatariinaParnanen/MobileGeneticElementDatabase, downloaded on 2020-06-28) consisting of 2714 unique MGE sequences including transposases; integrons of classes 1, 2, and 3; the integron associated disinfectant resistance gene *qacEΔ;* and IS and ISCR type insertion sequences[49]. Class 1 integrons were studied using mapping results obtained for the *intI* gene, which encodes an integron integrase (*IntI*). Taxonomic profiling was performed using both Metaphlan3[50] and Metaxa2[51]. The counts for bacterial 16S rRNA from Metaxa2 were used for normalization of ARG and MGE counts. The lengths of the functional genes and the 16S rRNA SSU gene were taken into account in the normalization.

The accuracy of the normalization method was validated by comparing two distinct methods using the 16S rRNA and *rpoB* housekeeping genes. The *rpoB* counts were obtained using the translated sequence reads described above as input for hmmsearch[52] against a *rpoB* specific HMM profile (pf04563) from the Pfam-A database[53]. Hits for forward reads (R1) were counted. The correlation between the ARGs normalized to both 16S rRNA and *rpoB* was significant ($R^2 = 0.61$, $p = 4.8 \times 10^{-8}$), and we proceeded to use 16S rRNA genes for normalization.

To study plasmid-mediated colistin resistance genes (*mcr* genes) and their genetic environments for their theoretical mobility potential, a subset of the samples was assembled using MEGAHIT (v1.2.8)[54], including the samples BH02, BH48, BFH19, BFH26, BFH41, FH1, and FH9. The minimum contig length was set to 1,000 nt. Contigs harboring plasmid-mediated colistin resistance genes were identified using a BLAST[55] search against *mcr*

genes in the ResFinder database (n=54). Similarly, contigs containing class 1 integrons were identified and screened for putative ARGs located in the same contig.

Visualizing the genetic environments of contigs harboring *mcr-5.1* colistin ARGs was performed using the anvi'o (v7)[56] platform tool 'anvi-interactive'. For this, single profiles were formed for the samples BH02, BH48, BFH26, BFH41, and FH9 with the option '-cluster-contigs'. Then, a BLAST nr search of the contigs against the ResFinder and MobileGeneticElementDatabase was performed. The resulting annotations were imported into the individual sample profiles containing the contigs database and open reading frames predicted by Prodigal (v2.6.3)[57] with the command 'anvi-import-functions'. The other predicted genes, sharing the same genetic environment with the *mcr-5.1* gene, were identified by inspecting the resulting annotations from the BLAST nr search on the web BLAST 'Graphics' view.

To visualize the co-localization of multiple ARGs with Bandage[58], the assembled contigs were converted into assembly graphs using the MEGAHIT command 'megahit_toolkit contig2fastg'. The Bandage program was run with a BLAST[55] query of a custom database consisting of ARGs present in the ResFinder database and the integron sequence-related genes *qacEΔ* and *intI1*.

2.4. Statistical analysis

Statistical analyses were performed in R (v3.6.3), and the results were visualized using ggplot2 (v3.3.3)[59], pheatmap (v 1.0.12)[60], and patchwork (v1.1.0.9000)[61]. Vector maps were drawn using the package rnaturalearth (v0.1.0)[62]. All statistical analyses described below were performed for all 67 HWW samples described in the 'Sample details' section and Supplementary Material Table S1B.

To model the total relative abundances of ARGs and MGEs in HWWs from Benin, Burkina Faso, and Finland, a generalized linear model (GLM) with gamma distribution and logarithmic link was fitted to 16S rRNA normalized count data. The effects of the country, hospital, and hospital section on the abundance of ARGs and MGEs were studied with the function 'glm' in the package MASS (v7.3-53.1)[63].

The 'ordinate' function in the Phyloseq package (v1.30.0)[64] was used for principal coordinate analysis (PCoA) of resistomes, mobilomes, and taxonomical compositions. The significance of the distinct clustering between the HWWs from different countries was tested using pair-wise adonis in the package vegan (v2.5-7)[65] with equal sample groups. For this, eight samples per country were selected randomly among the HWWs. Unique and shared ARGs were visualized with the function 'draw.triple.venn' in the package VennDiagram (v1.6.20)[66].

Pearson correlation coefficients and adjusted *p*-values for each ARG for class 1 integron and *qacEΔ* genes were obtained with the function 'corr.test' in the package psych (v2.1.3)[67] and visualized using packege ggcorrplot (v0.1.3)[68]. The correlation tests for different normalization methods (*rpoB* and 16S rRNA) were performed similarly.

Country-related differences in the prevalence of specific resistance determinants were examined using DESeq2 (v1.26.0)[69]. ARGs filtered using a variance threshold of 50 were screened for significantly ($p < 0.05$) different relative abundances represented by log-2-fold changes across the HWWs of the three countries in a pair-wise manner (Benin-Finland, Burkina Faso-Finland). The DESeq2 algorithm is designed for unnormalized count data from gene expression studies. Thus, to take into account the amount of bacterial DNA in the samples (counts of detected 16S rRNA), the relative abundances were multiplied by 10^5 and a pseudo count of one was added to avoid zero values unsuitable for integer transformation. A similar analysis using DESeq2 was performed to study differentially abundant taxa found in the HWWs from Benin or Burkina Faso versus Finland. The relative abundances obtained from Metaphlan3[50] were multiplied with the corresponding 16S rRNA counts for each sample from the Metaxa2[51] output. Analyses for differentially abundant genera and species were conducted separately.

3. Results

3.1. Features of hospital wastewaters (HWWs) in Benin, Burkina Faso, and Finland

We analyzed hospital wastewater-associated resistomes from Benin, Burkina Faso, and Finland. On average, 31 million sequence reads per sample were analyzed. Our definition of ARGs followed the scheme used by the ResFinder database, which relies on confirmed association between ARGs and their resistance phenotype[47].

HWWs from Benin showed the highest and from Burkina Faso the second-highest abundance of all detected ARGs normalized to bacterial 16S rRNA genes, here referred to as the relative sum abundance of ARGs. The lowest relative sum abundance was detected in HWWs from Finland (Figure 1a, Supplementary Material, Table S2A). No significant differences were observed in the relative ARG sums among the different hospitals in each country or between the hospital sections or wards within the sampled hospitals (Supplementary Material, Table S2A).

There were no significant differences in the relative sum abundance of all MGEs between the HWWs of the three countries (Figure 1b, Supplementary Material, Table S2A). Instead, country-wise differences for the gene *intI1* of class 1 integrons (Figure 1c) followed the trend observed for relative ARG abundances (Figure 1a). The observed high Pearson correlation of 0.80 ($p = 3.2 \times 10^{-16}$) between ARGs and *intI1* genes across all 67 HWW samples indicated an association between the genes.

In total, 2,331 different ARGs were detected in HWWs from Benin, Burkina Faso, and Finland (Supplementary Results, Figure S4). Of these, 1,280 were observed in at least one sample from every country. The resistomes of HWWs from Benin and Burkina Faso resembled each other more than those of the Finnish HWWs (Figure 2a). The resistomes in Benin and Burkina Faso shared 384 ARGs, while Finland and Benin shared only 15 ARGs and Finland and Burkina Faso shared 134 ARGs (Supplementary Results, Figure S4). The resistome compositions differed the most between the Beninese and Finnish HWW samples (Figure 2a). The country explained the differences in the resistomes by 29.5–33.7 % in the Benin-Finland comparison, while for the Burkina Faso-Finland

pair, the country explained 24.5–25.6 % (Supplementary Results, Table 1). In addition, while the ARGs found uniquely in HWWs from Burkina Faso made up approximately 14 % of all the detected ARGs in this study, the corresponding percentage for HWWs from Benin was only 2.5 % (Supplementary Results, Figure S4).

As the resistomes of West African countries and Finland were divergent, the drivers for the differences were subsequently investigated. The country explained less of the variance between taxonomical compositions than resistomes (Figure 2b, Supplementary Results, Table S1). A set of samples from a single hospital (Hospital 1, Supplementary Results, Figure S3b) seemed to drive the distinct clustering seen in the microbial community in HWWs from Burkina Faso.

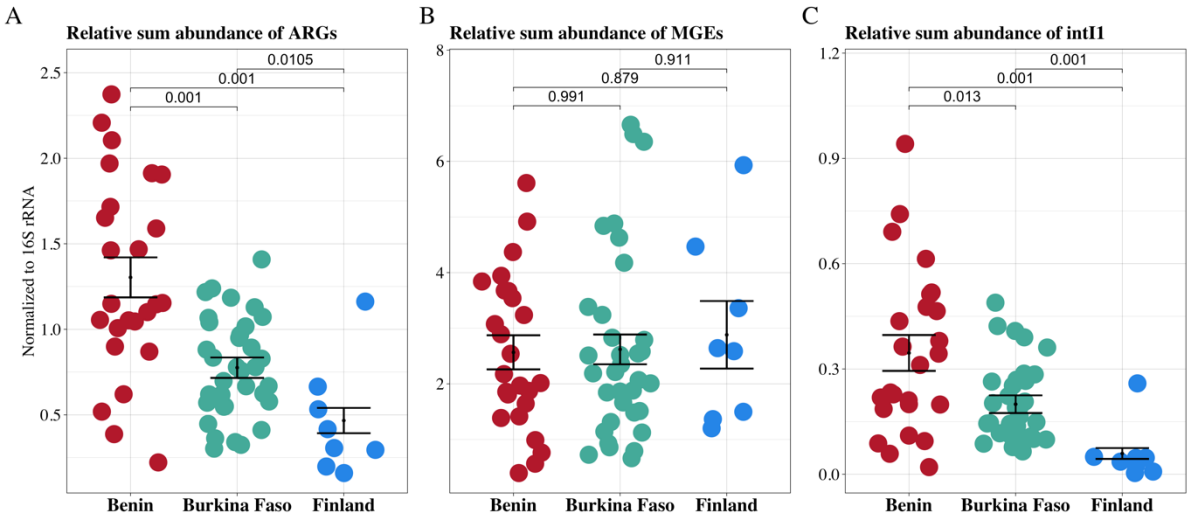


Figure 1: Relative sum abundance of ARGs, MGEs, and class 1 integrons. A generalized linear model (GLM) was fitted to 16S rRNA, and gene length normalized **a)** ARG **b)** MGE, and **c)** *intI1* count data to model the differences in the levels of these genes in the three countries. The vertical line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

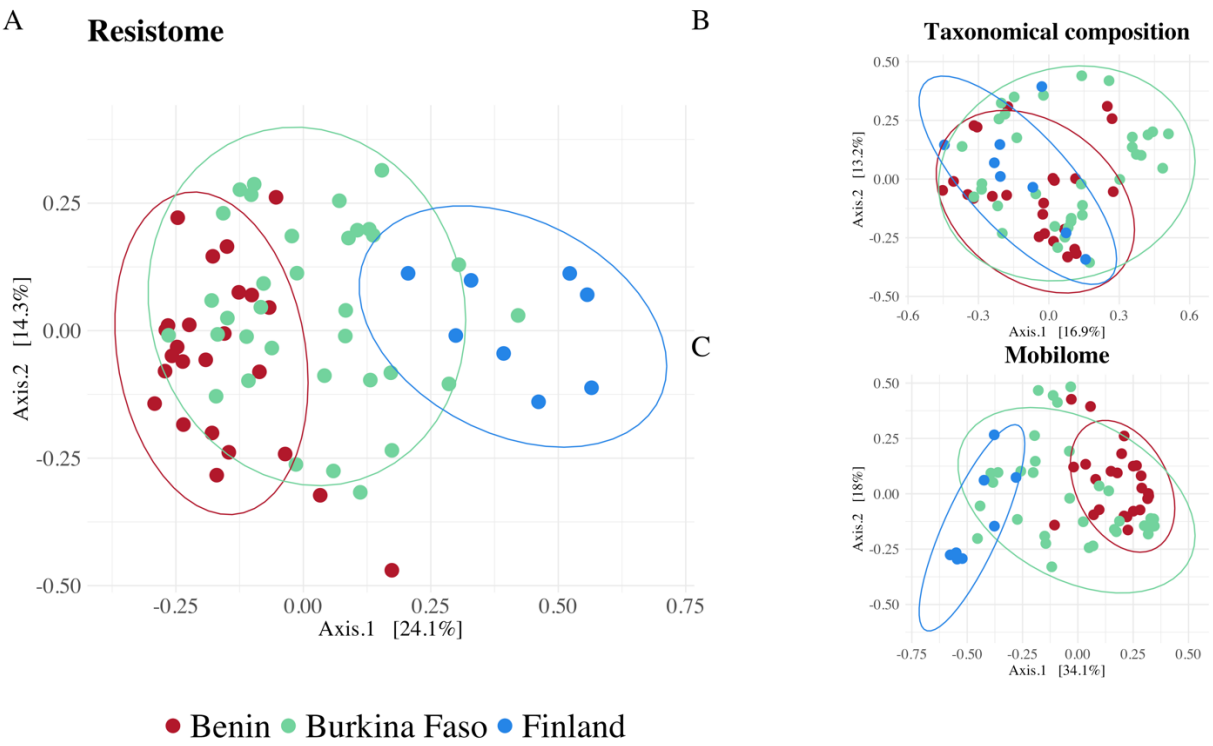


Figure 2: **PCoA plots describing the ordinations based on a) resistomes, b) taxonomical compositions, and c) mobilomes present in HWWs from Benin, Burkina Faso, and Finland.** All 67 HWW samples were included in the plots. However, due to unequal group sizes, groups of random sample sets were applied for testing the significance of the clustering based on the country. The tests were performed in triplicate with different samples to account for variation caused by sample selection (Supplementary Results, Table 1). Confidence ellipses represent 90 % confidence levels.

3.2. Clinically relevant taxa and associated ARGs

Taxonomical composition present in septic tanks and sumps in Benin and Burkina Faso differed from those found in hospital sewers in Finland. In Benin and Burkina Faso, the 15 most abundant genera present in HWW included *Desulfobulbus, Sulfurospirillum, Thauera*, and *Geobacter*, while in Finnish HWW, *Bifidobacterium* and *Faecalibacterium* were among the most abundant genera (Supplementary Material, Table S3). Opportunistic pathogens such as *Escherichia* and *Aeromonas* were detected among the 15 most abundant genera in the HWWs of all three countries (Figure 3, Supplementary Material, Table S3). *Klebsiella pneumoniae* was among the 15 most abundant species in Benin and Burkina Faso in terms of relative abundance (Figure 3, Supplementary Material, Table S3). In contrast, at the species level, *Escherichia coli, Enterococcus faecium*, and species of *Enterobacter cloacae* and *Pseudomonas aeruginosa* complexes were present but rare in all the countries. *Staphylococcus aureus* was detected exclusively in Burkinabe samples (Figure 3). Other species in the genus *Staphylococcus* were more commonly detected in HWWs from Benin and Burkina Faso compared to Finland (Supplementary Material, Table S4B and S5B).

Beta-lactamase genes of the *mec* operon, typically associated to methicillin-resistant *S. aureus* (MRSA)[17] but sometimes also found in other *Staphylococcus* species[70], were detected in HWWs from Benin (hospitals A, B & C) and Burkina Faso (hospitals I & J), but not from Finland (Supplementary Material, Tables S8, S9, and S10).

In contrast, vancomycin-resistant *Enterococci* (VRE) associated ARGs of the *van* operon were detected more commonly in HWWs from Finland compared to Benin and Burkina Faso (Figure 5, Supplementary Material, Tables S7, and S8). Several of these genes, such as *vanHBX*, *vanHDX,* and *vanC3XY,* were also unique to the Finnish samples (Supplementary Material, Table S10).

There was no significant difference (adjusted *p*-value < 0.05) in the abundance of the genus *Acinetobacter* between the HWWs from the three countries, but the abundances of different species of *Acinetobacter* varied significantly. *A. johnsonii* and carbapenemases in the *bla*OXA-211-like group (for example, *bla*OXA-211, -212, -280, and -281) that are intrinsically carried by this species were significantly more abundant in HWWs from Finland compared to either Benin or Burkina Faso (Figure 3 and Figure 4, Supplementary Material, Tables S4A, S5A, S6, and S7). Certain species of *Acinetobacter* that are not intrinsically carbapenem-resistant, such as *A. townerii*, were detected at significantly higher relative abundances in HWWs from Benin and Burkina Faso compared to Finland (Supplementary Material, Tables S4A, and S5A).

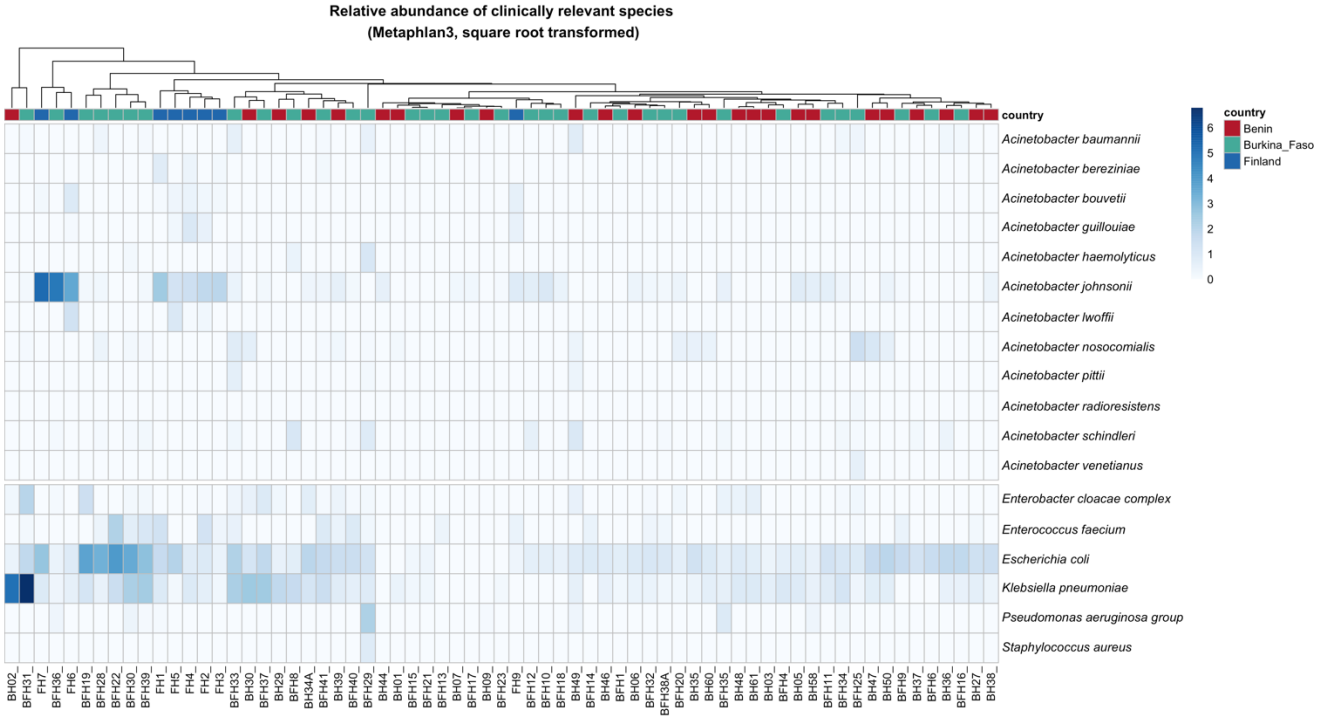


Figure 3: Relative abundance of clinically relevant species. The heatmap shows the relative abundances of species listed in the ESKAPEEc group (*Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa, Enterobacter* spp., *and Escherichia coli*) defined by de Angelis and colleagues[15]. In addition, *Acinetobacter* species with known intrinsic resistance to carbapenems according to the Beta-lactamase Database[71] are shown. Of these, species with a total relative abundance higher than 0.05 in the samples are shown.

3.3. Carbapenem resistance

The presence of seven acquired carbapenemases (*bla*GES, *bla*IMP, *bla*KPC, *bla*NDM, *bla*OXA-48, *bla*OXA-58, and *bla*VIM) was analyzed in detail due to their ability to cause complex infections with minimal treatment options when in a relevant host (Figure 4a). In HWWs from Benin, the carbapenemase genes in the *bla*GES family seemed

to dominate over other carbapenemases in all four hospitals. Most of the other carbapenemases were present at similar prevalence levels in HWWs from Burkina Faso and Finland (Figure 4a). This observation is in line with the lower number of different ARGs in Benin versus Burkina Faso (Supplementary Results, Figure S4).

Among the studied acquired oxacillinases, *bla*OXA-48 was detected solely in two samples in Burkina Faso. Instead, *bla*OXA-58 was present in the majority of Burkinabe and Finnish samples (Figure 4a). Most *bla*KPC genes were exclusively detected in HWWs from Finland (Figure 4a, Supplementary Material, Table S10). *bla*VIM, *bla*IMP, and *bla*NDM were detected in Burkinabe and Finnish HWW samples (Figure 4a), but they were more abundant in HWWs from Burkina Faso.

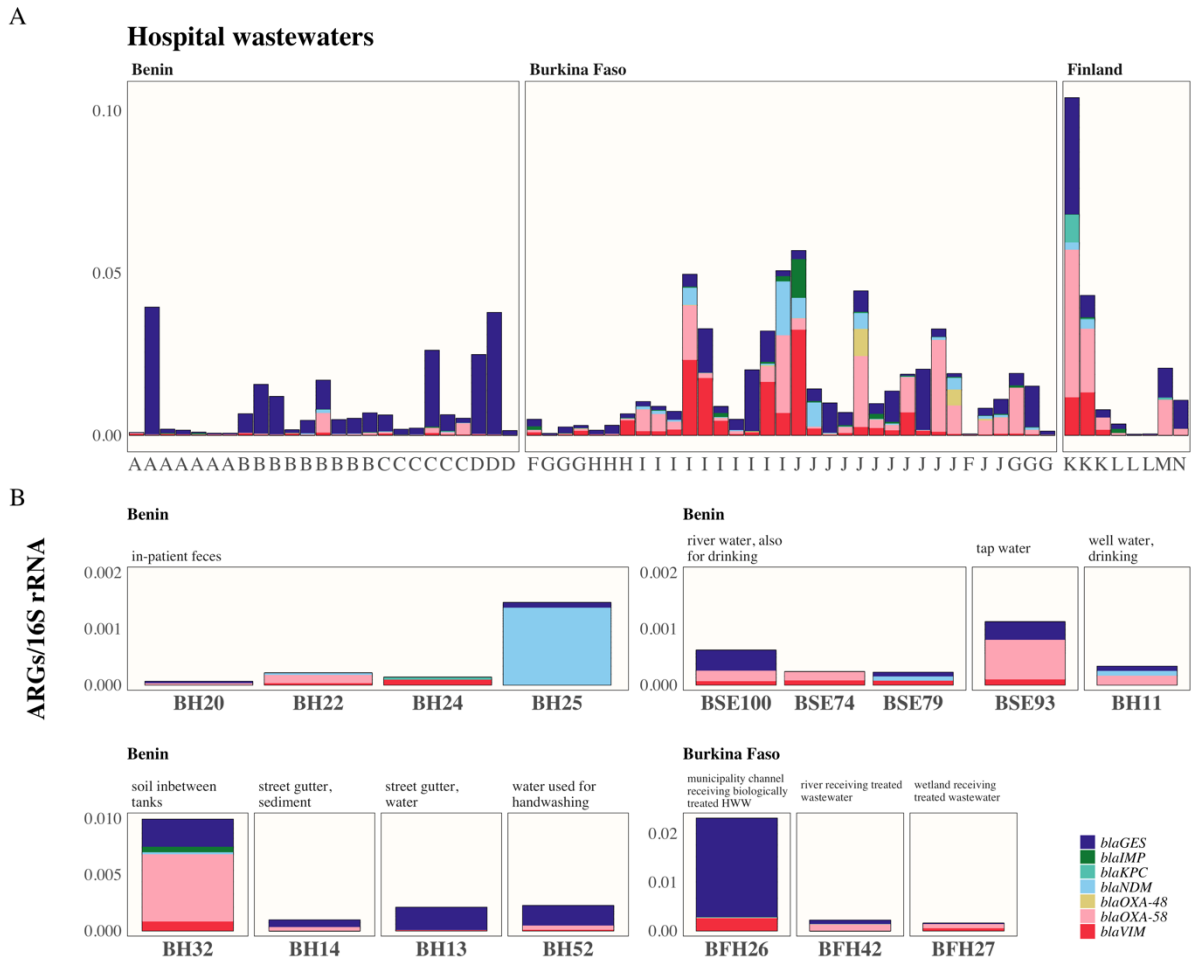


Figure 4: Abundance of carbapenemase genes in relation to the abundance of 16S rRNA genes in a) HWWs from Benin, Burkina Faso, and Finland, and b) various other sample types in Benin and Burkina Faso. Only the *bla*GES genes that encode carbapenemase production according to the Beta-lactamase Database[71] are shown, omitting the *bla*GES genes that encode ESBL. All screened carbapenemase gene variants are listed in Supplementary Material, Table S16.

3.4. Putative role of MGEs and class 1 integrons in shaping HWW resistomes

Three methods were used, each with its limitations, to study whether the abundance of class 1 integrons could explain the differences in the relative sum abundances of ARGs between the HWWs of the three countries (as

371 shown in Figure 1a). First, correlations between the relative abundances of the detected ARGs and the two class
372 1 integron-associated genes, *intI1* and *qacEΔ*, were studied. These correlations were visualized for those ARGs
373 differentially abundant in HWWs from each country in pair-wise comparisons (Figure 5). Second, contigs carrying
374 the *intI1* gene were studied for the putative co-localization of ARG sequences (Supplementary Material, Tables
375 S11A, and S11B). Third, fragments carrying multiple acquired ARGs were screened visually using assembly
376 graphs (Figure 6).
377
378 The top 20 ARGs that were more abundant in the Beninese HWWs showed stronger positive correlations with
379 class 1 integron genes compared to the ARGs that were more abundant in Burkina Faso and Finland (Figure 5).
380 This suggests a stronger class I integron association of ARGs in Beninese HWWs than in the other two countries.
381 The detection of some of these ARGs in contigs carrying several other acquired ARGs (Figure 6) or class 1
382 integron-associated genes (Supplementary Material, Table S12) suggests that the genes are linked to class 1
383 integrons. For example, the plasmid-mediated ARGs *qnrVC4* and *qnrVC5* were strongly positively correlated
384 with *intI1* and *qacEΔ*. The genes were located on contigs with other ARGs such as aminoglycoside ARGs and
385 either carbapenemase (Figure 6a) or ESBL genes in the *bla*GES gene family (Figure 6b, Supplementary Material
386 Table S6). While *qnrVC4* and *qnrVC5* were more abundant in HWWs from Benin and Burkina Faso than in those
387 from Finland, quinolone ARGs unassociated with integrons, such as *qnrB* variants, were instead more common
388 in HWWs from Finland (Figure 5, Supplementary Material Table S6, and S7).
389
390 Furthermore, the *bla*VEB ESBL genes and carbenicillinase beta-lactamases such as *bla*CARB and *bla*AER were
391 positively correlated with class 1 integrons in Benin and Burkina Faso (Figure 5, Supplementary Material Tables
392 S6, S7, S11A, and S11B). In contrast, *bla*MOX ESBL ARGs, which had no significant correlation with *intI1* or
393 *qacEΔ,* were significantly more abundant in HWWs from Finland compared to Benin or Burkina Faso (Figure 5,
394 Supplementary Material Tables S6, S7, S11A, S11B, and S14).
395
396 Oxacillinases constituted half of the top 20 most abundant ARGs in HWWs from Finland (Figure 5). Most of
397 these oxacillinases (*bla*OXA-211, -212, -334, -373, -309, -280, and -281) belong to the group of *bla*OXA-211-
398 like carbapenemases typically intrinsically carried by certain *Acinetobacter* species and showed no significant
399 correlation with *intiI1* or *qacEΔ* (Figure 5, Supplementary Material Tables S6, S7, S11A, and S11B). In contrast,
400 certain genes of oxacillinases from other *bla*OXA gene families, which were significantly more abundant in
401 HWWs from Benin or Burkina Faso, showed significant correlation with *intI1* and *qacEΔ* (Figure 5) and were
402 found in shared contigs with the *intI1* gene (Supplementary Material, Table S12) and other ARGs (Figure 6c).

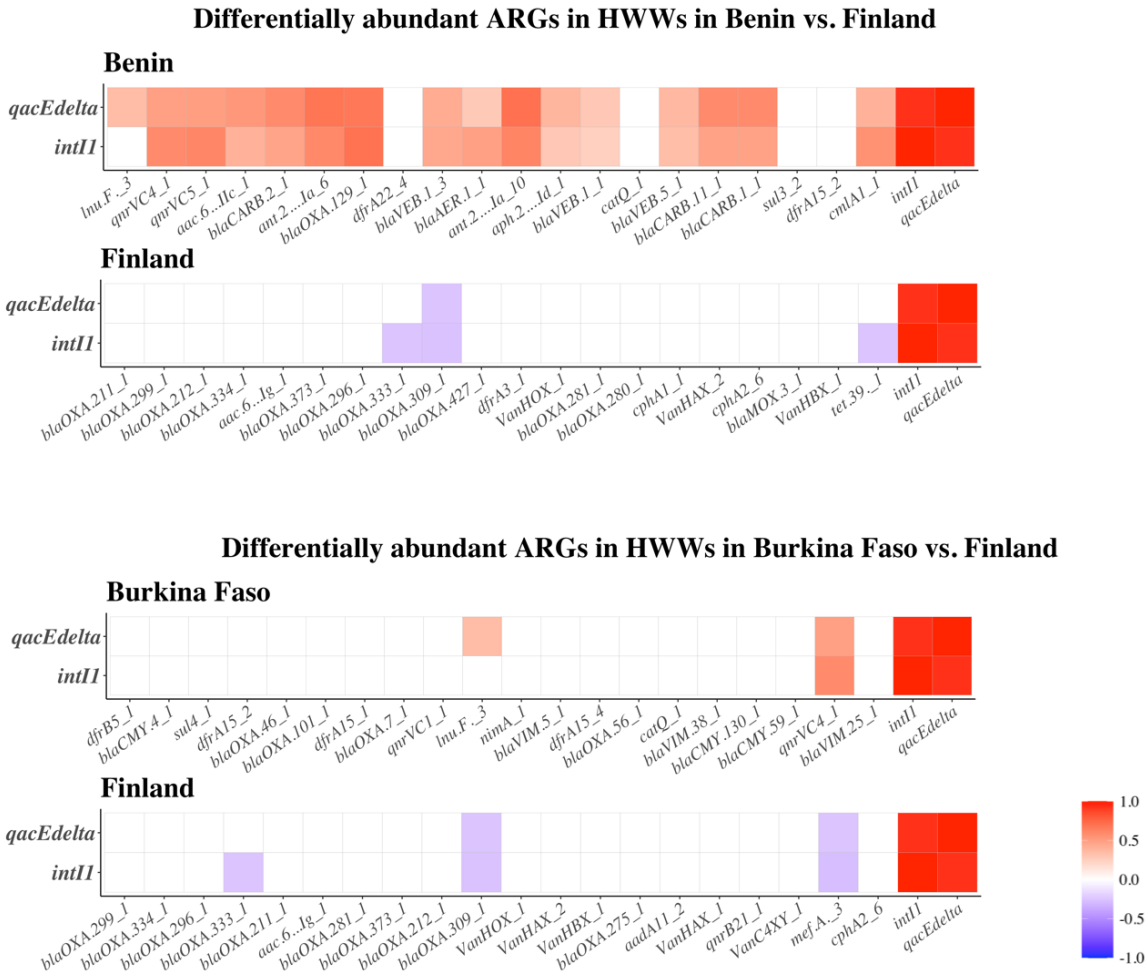
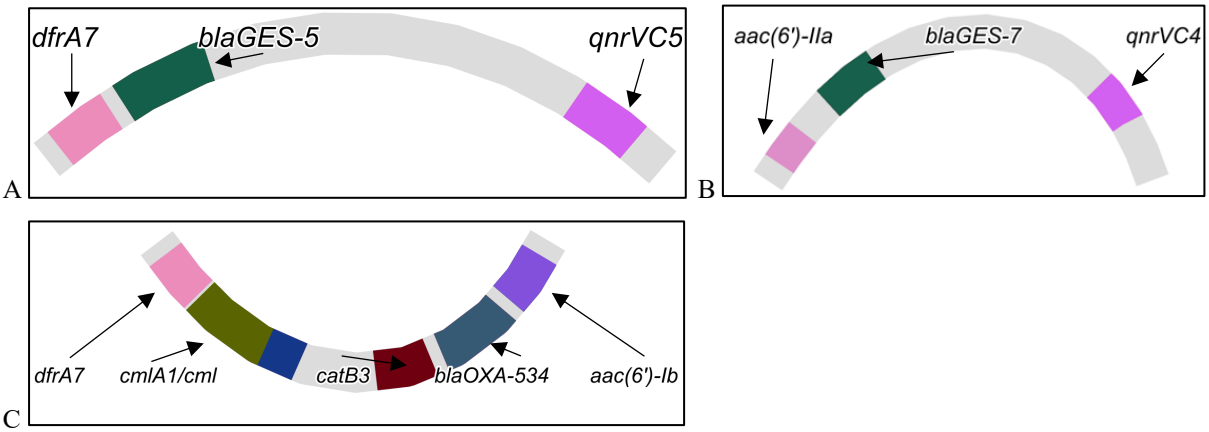


Figure 5: Differently abundant ARGs in HWWs from Benin and Burkina Faso and their putative associations with class 1 integrons. 20 significantly more abundant ARGs in HWWs of each country in pair-wise comparisons (Benin-Finland, Burkina Faso-Finland) are shown. Pearson correlations between each ARG and the genes *intI1* and *qacEΔ* associated with class 1 integrons (Supplementary Material, Table S11A, and Table S11B) are denoted with color coding. The red color indicates positive and blue negative correlation. Non-significant (adjusted $p > 0.05$) correlations are marked as blank (white) squares.



414 **Figure 6. Examples of assembly graphs showing multidrug-resistance fragments putatively encoded by class 1 integron gene cassettes**
415 **in samples a) BH02, b) BH48, and c) BFH19.**
416
417 *3.5. Plasmid-mediated colistin resistance genes*
418
419 Mobile colistin resistance genes (*mcr* genes) were observed in several HWW samples in Benin, Burkina Faso,
420 and Finland (Table 1). Variants of *mcr-5* (genes *mcr-5.1* and *mcr-5.2*) were the most common as they were found
421 in HWWs from all except one hospital (hospital F) in the three countries. We were able to confirm the location of
422 *mcr-5* within a Tn3-like transposase element in the metagenome-assembled contigs (Supplementary Results,
423 Figure S5, Supplementary Material, Table S13).
424
425 **Table 1: Detection of different plasmid-mediated colistin resistance genes (*mcr* genes).** The *mcr* genes genes found in ResFinder database[47]
426 were clustered based on 90 % similarity in their sequence identity defined using CD-HIT (v4.8.1)[72].

	Average relative abundance				Positive samples			
Gene variant	**Benin**	**Burkina Faso**	**Finland**	**All**	**Benin**	**Burkina Faso**	**Finland**	**All**
mcr-1	0	4.61×10^{-7}	1.97×10^{-6}	4.69×10^{-7}	0 / 25	1 / 34	1 / 8	2 / 67
mcr-3	8.5×10^{-4}	4.18×10^{-4}	1.58×10^{-4}	5.49×10^{-4}	16 / 25	26 / 34	7 / 8	42 / 67
mcr-3.17	7.29×10^{-6}	3.74×10^{-6}	1.32×10^{-4}	2.04×10^{-5}	3 / 25	6 / 34	5 / 8	14 / 67
mcr-4	3.56×10^{-6}	2.63×10^{-6}	0	2.66×10^{-6}	1 / 25	3 / 34	0 / 8	4 / 67
mcr-5	9.98×10^{-4}	1.02×10^{-3}	7.53×10^{-4}	9.79×10^{-4}	22 / 25	26 / 34	8 / 8	52 / 67
mcr-7	5.54×10^{-5}	1.54×10^{-4}	1.94×10^{-4}	1.22×10^{-4}	10 / 25	23 / 34	8 / 8	39 / 67
mcr-8	0	0	1.68×10^{-6}	2.01×10^{-7}	0 / 25	0 / 34	1 / 8	1 / 67
mcr-9	1.08×10^{-6}	2.14×10^{-6}	3.08×10^{-4}	3.83×10^{-5}	1 / 25	6 / 34	4 / 8	10 / 67
mcr-10	1.22×10^{-4}	1.03×10^{-3}	1.97×10^{-5}	5.69×10^{-4}	11 / 25	20 / 34	2 / 8	33 / 67

427
428
429 In addition to the HWW samples, high relative abundances of *mcr-5* genes were also detected in the immediate
430 and more distant surroundings of the hospitals, such as in the street gutter water 100 m away from hospital B and
431 in the soil between two HWW septic tanks in Benin (Supplementary Results, Table 2). For hospital C, the relative
432 abundance of *mcr-5* in a puddle surrounding the surgery room HWW septic tank was greater than the average
433 relative sum abundance of *mcr-5* in the actual HWW septic tanks of that hospital (Supplementary Results, Figure
434 S2c and Table S2).
435
436 In Burkina Faso, very high relative sum abundances of *mcr-5* were observed in a municipality channel leading
437 out from hospital I after passing through a biological treatment process. Furthermore, the presence of *mcr*-5 was
438 also detected in nearby wetland where the processed wastewater is released (Supplementary Results, Table S2).
439 Similarly, in another Burkinabe city, *mcr-5* was found in a sample before wastewater treatment and in the river
440 receiving the water after treatment (Supplementary Results, Table S2).
441
442 The next most prevalent *mcr*-genes, namely *mcr-3, mcr-10,* and *mcr-7*, were detected in certain cases at very high
443 relative abundances in the HWW samples (Table 1). The gene *mcr-3* was also present in the water used for

handwashing at hospital C (Supplementary Results, Figure S2e, Supplementary Results, Table S2). The gene *mcr-2* was not detected, and the lowest possible average relative sum abundance was detected for the gene *mcr-1*.

3.6. Other than hospital wastewater samples

Other sample types within and outside the studied hospitals were investigated to obtain a more comprehensive understanding of ARG prevalence. Metronidazole resistance genes, such as *nimA* and *nimE,* were among the top 15 most abundant ARGs in the fecal samples of the four in-patients at a Beninese hospital B. (Supplementary Material, Table S14). These ARGs were also significantly more abundant in HWWs from Benin and Burkina Faso compared to Finland (Figure 5, Supplementary Material, Tables S6 and S7). *bla*NDM was detected in one of the four fecal samples (Figure 4b) and was also present in the HWWs of hospital B (Figure 4b and 4a). Other high abundance ARGs in the fecal samples included ESBL and broad-spectrum beta-lactamase genes of the *cfxA* family (Supplementary Material, Table S14).

As in the HWWs from Benin (Figure 4a), *bla*GES carbapenemases were also prevalent in other sample types from a hospital environment, such as in the water intended for handwashing (Supplementary Results, Figure S2e) for hospital C, soil between HWW septic tanks in the yard of hospital B, and street gutter water 100 meters away from hospital B (Figure 4d). *bla*GES carbapenemase genes were also found in various waters used for drinking, including river waters in a remote countryside village (Figure 4c). However, the relative sum abundance of ARGs was low in the three rivers in central Benin (Figure 4c). We also detected *bla*NDM in well water used for drinking in Benin (Figure 4c). Albeit at low relative abundances, *bla*GES, *bla*VIM, and *bla*OXA-58-like carbapenemases were detected in natural waters receiving treated HWWs from Burkina Faso (Figure 4d). Significantly higher abundances of these ARGs were found in biologically treated wastewater before release into the river (Figure 4c).

4. Discussion

We characterized the bacterial community composition and resistome in 67 HWWs and 16 related samples from Benin and Burkina Faso and compared them to HWWs from Finland. This is the first study investigating hospital wastewaters in the two West African countries using a metagenomic approach. Due to the lack of systematic AMR surveillance in these countries, available data on bacterial resistance are patchy and heterogenous[1,4,31,34,37]. Thus, the magnitude of the resistance problem and the specific reservoirs of ARGs in many African countries are yet to be unraveled. The relative sum abundance of all ARGs detected in HWWs was the highest in Benin, yet the number of different ARGs detected there was the lowest. The differences in resistomes between the countries seemed to be only partially explained by variation in taxonomical composition, as the clustering of taxonomical composition was not as strongly correlated by country compared to clustering of resistomes. This suggests that despite differences in the HWW collection systems in the studied countries (HWW is relatively stagnant in septic tanks or sumps in Benin and Burkina Faso and flowing in HWW drains in Finland), the microbial community compositions present appear to be similar. However, among the most abundant taxa, we did find many strictly environmental bacteria in HWWs from Benin and Burkina Faso but not in Finland. This may be due to the closer contact of water with environmental surfaces in septic tanks and sumps compared to the more isolated wastewater

484 systems in Finland. Bacterial genera typically found in soil and aquatic environments[73] have also been previously
485 described to dominate HWWs from Benin and Burkina Faso[74]. In contrast, the 15 most abundant genera in Finnish
486 HWWs included less environmental bacteria and more bacteria typical for human gut communities, such as
487 *Bifidobacterium*[49] and *Faecalibacterium*[75]. Another factor possibly explaining the differences in the dominant
488 taxa present in the HWWs from Benin and Burkina Faso in comparison to Finland could arise from the fact that
489 in Finland, the hospital toilet waters containing human fecal material are directed to HWWs, while for the majority
490 of the West African HWWs studied here, this was not the case.
491
492 The ARGs that were more abundant in HWWs from Benin compared to Finland appeared to be more commonly
493 carried by class 1 integrons. They were strongly correlated with class 1 integron-associated genes (*intI1* and
494 *qacEΔ*[10]) and, in certain cases, detected in shared contigs with these genes. Moreover, we could detect gene
495 cassettes consisting of several different acquired ARGs linked to each other, likely referring to gene cassettes
496 carried by class 1 integrons[11]. Furthermore, a higher abundance of ARGs intrinsically carried by specific species,
497 such as *bla*OXA-211 genes by *A. johnsonii*[76,77], was a distinctive feature of the HWW resistomes from Finland.
498 Intrinsic resistance is often chromosomally encoded and rarely mobile or related to horizontal gene transfer[78],
499 whereas acquired resistance is typically carried by MGEs, such as class 1 integrons. As expected, no significant
500 correlations between these intrinsic *bla*OXA genes and class 1 integron associated genes (*intI1* and *qacEΔ*[10]) were
501 found, supporting the observation that ARGs putatively carried by class 1 integrons had a smaller impact on
502 resistomes in HWWs from Finland compared to Benin. Based on these results, we speculate that the high sum
503 abundance of ARGs seen in HWWs from Benin may, in part, be explained by the mobilome, more specifically,
504 the prevalence of class 1 integrons. However, correlation does not necessarily indicate causation, and in this study,
505 quantitative inferences cannot be made for metagenome-assembled contigs and the putative co-localization of
506 ARGs and class 1 integron-associated genes (*intI1* and *qacEΔ*[10]). MGEs other than class 1 integron associated
507 genes did not exhibit significant correlation with ARGs. This is in line with the current understanding of the role
508 of class 1 integrons, in particular, among all MGEs in spreading acquired ARGs[9].
509
510 Interestingly, the ARGs observed in HWWs from Benin were the highest in number but probably lower in clinical
511 importance compared to those in the HWWs from Burkina Faso and Finland. While the most significant
512 carbapenemase genes *bla*KPC, *bla*NDM, *bla*VIM, *bla*IMP, *bla*OXA-48, and *bla*OXA-58 were found in several
513 HWW samples from Burkina Faso and Finland at varying abundances, the abundance of many of them was low,
514 or they were not detected at all in HWWs from Benin. The *bla*GES carbapenemase genes were an exception. They
515 were the most prevalent of the seven carbapenemases in Beninese HWWs and were also found in samples
516 unrelated to hospitals in Benin.
517
518 We suspect that as *bla*GES carbapenemases are typically carried by class 1 integrons, as confirmed by previous
519 studies[79], a co-selection phenomenon[3,80] might play a role in increasing their spread also in here. In this scenario,
520 residues of antibiotics other than carbapenems would cause co-selection of *bla*GES. We also found a high relative
521 abundance of plasmid-mediated quinolone ARGs in the HWWs from Benin. These *qnrVC* genes showed strong
522 correlations to class 1 integron associated genes (*intI1* and *qacEΔ*[10]) and were in some cases found to be located
523 in the same contigs with *bla*GES-5 carbapenemase genes (Figure 6b). Hence, selection targeting quinolone ARGs

524 could co-select for carbapenem resistance encoded by the gene *bla*GES-5. Altogether our observations of ARGs
525 putatively carried by class 1 integrons are worrisome as new ARG combinations among integron gene cassettes
526 enable their selection by a broader set of antimicrobials or biocides such as quaternary ammonium compounds
527 (QACs)[10,81,82].
528
529 Despite the lowest total relative abundance of ARGs detected in HWWs from Finland, some of the seven
530 carbapenemase genes of special interest seemed to have even higher abundances in Finland than in the surveyed
531 West African countries. However, the occurrence of these ARGs was not homogenous among the Finnish samples.
532 For instance, a high relative abundance *bla*KPC was observed in only one of the Finnish samples. Furthermore,
533 the overall occurrence of *bla*KPC across all HWWs from Benin, Burkina Faso and Finland studied here aligned
534 with many local and global reports and systematic reviews[83–86], according to which *bla*KPC carbapenemases are
535 spreading more profusely in Europe and North America than in Africa. Similarly, the relative abundance of
536 *bla*OXA-58 was higher among Finnish than Beninese or Burkinabe HWW samples.
537
538 Disconcertingly, *bla*NDM was found in the fecal sample of an in-patient staying at a Beninese hospital. *bla*NDM
539 was also detected in the wastewaters of the hospital, indicating the sensitivity of wastewater-based monitoring of
540 bacterial resistance carried by hospital in-patients. *bla*NDM carried by *A. baumannii* has recently, in 2021, been
541 reported in clinical settings in Benin[87].
542
543 The *mec* operon beta-lactamase genes carried by some *Staphylococcus* strains were detected only in HWWs from
544 Benin and Burkina Faso and not in those from Finland. While a decline in hospital-acquired MRSA infections has
545 been observed in high- and middle-income countries (HMICs) due to improved hospital infection control and
546 alternative antibiotics[17,40], MRSA remains a high-risk resistant pathogen with elevated and underestimated
547 abundance in many LICs, including African countries[3,88,89]. Our observations of the incidence of *mec* genes seem
548 to align with this epidemiological data[17,40,90]. However, we acknowledge that *mec* genes can also be found in some
549 strains of other *Staphylococcus* species[70], and as we are not able to identify the carrier species of the *mec* genes
550 detected here, the presence of MRSA strains in these hospitals cannot be confirmed.
551
552 The more common occurrence of *van* genes in HWWs from Finland compared to Benin or Burkina Faso suggests
553 that vancomycin-resistant *Enterococci* (VRE) causing severe difficult-to-treat-infections[16] are more frequent in
554 hospitals in Finland, possibly due to differences in the consumption of vancomycin in different countries. Only
555 one of the West African hospitals reported vancomycin among the most used antibiotics (Supplementary Material,
556 Table S1A). In HMICs, the use of vancomycin is more common[91], and VRE is increasing in Europe[40].
557
558 Our observation of the high prevalence of metronidazole ARGs in Benin and Burkina Faso but not in Finland
559 most likely reflects differences in the consumption of this drug. In LICs, diarrheal infections caused by pathogens
560 such as *Clostridioides difficile* and other anaerobic bacteria, as well as protozoa, are common. The consumption
561 of metronidazole commonly used to treat these infections may be high in these countries as these drugs are being
562 prescribed, for example, both with and without a positive diagnostic test for *Clostridioides difficile*[92]. However,

563 multidrug-resistant strains of *C. difficile* seem to already be prevalent in hospitals in certain parts of Africa as a
564 consequence of the heavy use of antimicrobials among in-patients[92].
565
566 Compared to many HMICs such as those in Northern Europe, the use of antimicrobials in agricultural[93] and
567 clinical[4] settings is less controlled – or even entirely unregulated – in many African countries. For example, while
568 banned in many other countries, the use of colistin as a feed additive is allowed in many LICs[94], such as those in
569 Africa[93]. In contrast to previous reports on the prevalence of mobile colistin resistance globally[30] and Africa[31,95],
570 *mcr-5* was the most commonly detected gene in the HWWs from our study. Instead, our observations are in line
571 with the results obtained in Germany by Kneis and colleagues using both metagenomics and qPCR[96]. This
572 discrepancy may be due to differences between the methodologies used to screen for colistin resistance. Colistin
573 ARGs other than the best known genes *mcr-1, -2,* and *-3* are rarely targeted when screening for mobile colistin
574 resistance with conventional PCR[96]. Our study shows that to obtain a more realistic view of the colistin resistance
575 situation in Africa, screening should also be conducted for different *mcr* genes, as was recently done by Ngbede
576 and colleagues[97].
577
578 We confirmed that the *mcr-5* gene detected here was embedded in a Tn3-like transposase similar to previous
579 reports for *Salmonella enterica*[98,99] and *Escherichia coli*[100] plasmids and the chromosome of *Cupriavidus*
580 *gilardii*[98]. These Tn3-like elements are flanked by inverted repeats, which enable translocation and putatively a
581 wide host range for the *mcr-5.1* harboring element[98–100]. Based on our results, the occurrences of *mcr-5* and other
582 mobile colistin resistance genes in Benin and Burkina Faso were not restricted to HWW septic tanks or sumps.
583 For example, at a Beninese hospital, a water puddle in the hospital yard had a higher relative abundance of several
584 mobile colistin resistance genes, such as *mcr-5*, compared to the actual HWW septic tanks for that hospital.
585 Furthermore, different *mcr* genes, including *mcr-3* and *mcr-5,* were also detected in water intended for
586 handwashing of visitors to the hospital. Thus, the transmission of these clinically critical ARGs within hospitals
587 could be enabled via hands contaminated by this water. Similarly, Jacobs and colleagues have drawn attention to
588 the potentially inferior microbiological water quality in similar hand washing water tanks, which are very common
589 in West Africa[4].
590
591 In Burkina Faso, a high amount of colistin resistance genes was released into natural waters (2.36×10^{-4}) after
592 wastewater treatment. For comparison, a similar relative abundance of *mcr-5* genes (3.0×10^{-4}) was reported for
593 untreated municipal wastewater in Germany[96]. Although we cannot confirm the association between the *mcr-5*
594 genes detected in the natural waters and the hospital-associated wastewater in the Burkinabe hospital, our study
595 suggests the inability of the currently used wastewater treatment processes in Burkina Faso to remove colistin
596 resistance genes. Wastewater treatment systems have also earlier been described to be inadequate in Benin and
597 Burkina Faso[101]. This increases the risk of the further spread of colistin and other ARGs as inadequately treated
598 water is released into natural waters used for various purposes by local people. Furthermore, in many African
599 countries, untreated wastewater is used for irrigation in urban agriculture, possibly enabling the dissemination of
600 ARGs to fresh produce[102,103]. The situation appears to be especially critical in Cotonou, where the hospitals
601 involved in the study did not use any kind of wastewater treatment. Groundwater in the city is extracted from the
602 shallow aquifer, which is polluted due to unauthorized waste deposits, inadequate toilets, pit latrines, and septic

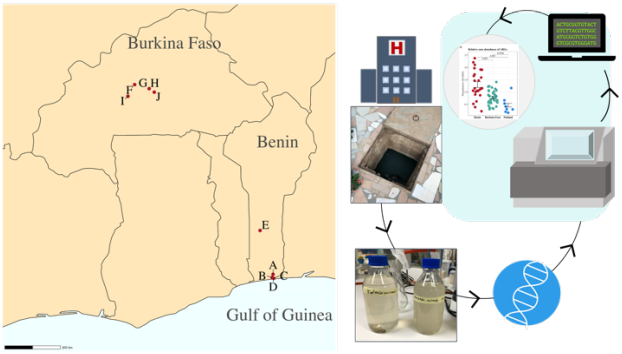
603 tanks prone to leakage and hydraulic failure[104]. Our study shows that untreated hospital wastewaters, which can
604 leak into groundwater and the environment, carry clinically dangerous antimicrobial-resistant bacteria and
605 resistance genes.
606
607 Taken together, based on our results from nine hospitals, we can state that clinically important ARGs are
608 circulating in West African hospitals and can spread to their surroundings. There was no significant difference in
609 ARG abundances based on the hospital section or different hospitals within these countries. Various carbapenem
610 ARGs were found in HWWs from Burkina Faso and Finland while among these *bla*GES seem to dominate in the
611 studied hospitals in Benin. Our results suggest that class 1 integron mediated co-selection could play an important
612 role in shaping the resistome in the Beninese samples. In addition to HWWs, carbapenemase genes and mobile
613 colistin resistance genes were detected in the vicinity as well as locations more distant from the hospitals. Finally,
614 there were significant differences between the HWW resistomes in Benin and Burkina Faso concerning the total
615 relative sum of ARGs and the prevalence of clinically relevant ARGs. Therefore, it is important to consider the
616 differences between African countries in AMR surveillance.
617
618 *Supporting Information*
619 The supporting figures and tables referred within the text are provided in separate Supplementary Results and
620 Supplementary Material files.
621
622 *Acknowledgments*
623 We thank the staff in all the participating hospitals for their collaboration in sampling. We also thank the laboratory
624 personnel of the URMAPha at the University of Abomey-Calavi in Benin and at the Clinical Research Unit of
625 Nanoro (CRUN) in Burkina Faso for sample processing. Finally, we acknowledge CSC, IT Center for Science,
626 Finland for providing the computational resources for the study.
627
628 *Funding Sources*
629 This work was supported by the Academy of Finland (Grant Numbers 318643 and 316708) and University of
630 Helsinki HiLife Grand Challenges funding.
631
632 *Ethical permission*
633 The ethical permissions for the project were received from "Comité National d´Ethique pour la Recherche en
634 Santé" under the Health Ministry in Benin and "Comité d´Ethique pour la Recherche en Santé" under the Health
635 Ministry in Burkina Faso.
636
637 *Data availability*
638 The data for this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under the
639 accession number PRJEB47975 and will be public upon article publication in a peer reviewed journal.
640
641 *Code availability*

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

1. Global Action Plan On Antimicrobial Resistance WHO Library Cataloguing-in-Publication Data Global Action Plan on Antimicrobial Resistance. (2015) doi:10.13140/RG.2.2.24460.77448.
2. Ventola, C. L. The Antibiotic Resistance Crisis Part 1: Causes and Threats. *Phys. Ther.* **40**, 277–83 (2015).
3. Laxminarayan, R. *et al.* Antibiotic resistance-the need for global solutions. *Lancet Infect. Dis.* **13**, 1057–1098 (2013).
4. Jacobs, J. *et al.* Diagnostic Bacteriology in District Hospitals in Sub-Saharan Africa: At the Forefront of the Containment of Antimicrobial Resistance. *Front. Med.* **6**, 205 (2019).
5. Hu, Y., Gao, G. F. & Zhu, B. The antibiotic resistome: gene flow in environments, animals and human beings. *Front. Med.* **11**, 161–168 (2017).
6. Tacconelli, E., Carrara, E., Savoldi, A., Kattula, D. & Burkert, F. Global Priority List Of Antibiotic-Resistant Bacteria To Guide Research, Discovery, And Development Of New Antibiotics. http://www.cdc.gov/drugresistance/threat-report-2013/ (2017).
7. Martínez, J. L., Coque, T. M. & Baquero, F. What is a resistance gene? Ranking risk in resistomes. *Nat. Rev. Microbiol.* **13**, 116–123 (2014).
8. Karkman, A., Berglund, F., Flach, C. F., Kristiansson, E. & Larsson, D. G. J. Predicting clinical resistance prevalence using sewage metagenomic data. *Commun. Biol.* **3**, 711 (2020).
9. Gillings, M. R. *et al.* Using the class 1 integron-integrase gene as a proxy for anthropogenic pollution. *ISME J.* **9**, 1269–1279 (2015).
10. Gillings, M. R. Integrons: Past, Present, and Future. *Microbiol. Mol. Biol. Rev.* **78**, 257–77 (2014).
11. Collis, C. M. & Hall, R. M. Expression of antibiotic resistance genes in the integrated cassettes of integrons. *Antimicrob. Agents Chemother.* **39**, 155–162 (1995).
12. Li, Y. *et al.* Molecular characterisation of antimicrobial resistance determinants and class 1 integrons of *Salmonella enterica* subsp. *enterica* serotype Enteritidis strains from retail food in China. *Food Control* **128**, 108191 (2021).
13. Gillings, M. *et al.* The Evolution of Class 1 Integrons and the Rise of Antibiotic Resistance. *J. Bacteriol.* **190**, 5095–5100 (2008).
14. Chen, B., Han, J., Dai, H. & Jia, P. Biocide-tolerance and antibiotic-resistance in community environments and risk of direct transfers to humans: Unintended consequences of community-wide surface disinfecting during COVID-19? *Environ. Pollut.* **283**, 117074 (2021).
15. De Angelis, G. *et al.* Incidence and antimicrobial resistance trends in bloodstream infections caused by ESKAPE and *Escherichia coli* at a large teaching hospital in Rome, a 9-year analysis (2007-2015). *Eur. J. Clin. Microbiol. Infect. Dis.* **37**, 1627–1636 (2007).
16. Alemayehu, T. & Hailemariam, M. Prevalence of vancomycin-resistant enterococcus in Africa in one health approach: a systematic review and meta-analysis. *Sci. Rep.* **10**, 20542 (2020).
17. Klein, E. Y., Sun, L., Smith, D. L. & Laxminarayan, R. The Changing Epidemiology of Methicillin-Resistant *Staphylococcus aureus* in the United States: A National Observational Study. *Am. J. Epidemiol.* **177**, 666–74 (2013).
18. Marturano, J. E., Ave, H., Marturano, J. E. & Lowery, T. J. ESKAPE Pathogens in Bloodstream Infections Are Associated With Higher Cost and Mortality but Can Be Predicted Using Diagnoses Upon Admission. *Open Forum Infect. Dis.* **6**, (2019).
19. Karkman, A., Do, T. T., Walsh, F. & Virta, M. P. J. Antibiotic-Resistance Genes in Waste Water. *Trends Microbiol.* **26**, 220–228 (2018).
20. Huijbers, P. M. C., Flach, C. F. & Larsson, D. G. J. A conceptual framework for the environmental surveillance of antibiotics and antibiotic resistance. *Environ. Int.* **130**, 104880 (2019).
21. World Health Organization. *Global Antimicrobial Resistance and Use Surveillance System (GLASS) Report 2021. Geneva*. (2021).
22. World Health Organization. WHO report on surveillance of antibiotic consumption: 2016-2018 early implementation. https://www.who.int/medicines/areas/rational_use/oms-amr-amc-report-2016-2018/en/ (2018).
23. Suay-García & Pérez-Gracia. Present and Future of Carbapenem-resistant Enterobacteriaceae (CRE) Infections. *Antibiotics* **8**, 122 (2019).
24. Poirel, L. *et al.* OXA-58, a novel class D β-lactamase involved in resistance to carbapenems in *Acinetobacter baumannii*. *Antimicrob. Agents Chemother.* **49**, 202–208 (2005).
25. Poirel, L. *et al.* GES-2, a class A β-lactamase from Pseudomonas aeruginosa with increased hydrolysis

of imipenem. *Antimicrob. Agents Chemother.* **45**, 2598–2603 (2001).
26. Nordmann, P. & Poirel, L. The difficult-to-control spread of carbapenemase producers among *Enterobacteriaceae* worldwide. *Clin. Microbiol. Infect.* **20**, 821–830 (2013).
27. Löfmark, S., Edlund, C. & Erik Nord, C. Metronidazole Is Still the Drug of Choice for Treatment of Anaerobic Infections. *Clin. Infect. Dis.* **1**, S16–S23 (2010).
28. Alauzet, C., Lozniewski, A. & El Ene Marchandin, H. Metronidazole resistance and *nim* genes in anaerobes: A review. *Anaerobe* **55**, 40–53 (2018).
29. Falagas, M. E. & Kasiakou Sofia. K. Colistin: The Revival of Polymyxins for the Management of Multidrug-Resistant Gram-Negative Bacterial Infections. *Clin. Infect. Dis.* **40**, 1333–41 (2005).
30. Ling, Z. *et al.* Epidemiology of mobile colistin resistance genes *mcr-1* to *mcr-9*. *J. Antimicrob. Chemother.* **75**, 3087–3095 (2020).
31. Anyanwu, M. U., Odilichukwu, C., Okpala, R., Chah, K. F. & Shoyinka, V. S. Prevalence and Traits of Mobile Colistin Resistance Gene Harbouring Isolates from Different Ecosystems in Africa. *Biomed Res. Int.* **22**, 6630379 (2021).
32. Collignon, P., Beggs, J. J., Walsh, T. R., Gandra, S. & Laxminarayan, R. Anthropological and socioeconomic factors contributing to global antimicrobial resistance: a univariate and multivariable analysis. *Lancet Planet Heal.* **2**, e398-405 (2018).
33. Van Boeckel, T. P. *et al.* Global trends in antimicrobial resistance in animals in low-and middle-income countries. *Science (80-.).* **365**, eaaw1944 (2019).
34. Iskandar, K. *et al.* Surveillance of antimicrobial resistance in low- and middle-income countries: a scattered picture. *Antimicrob Resist Infect Control* **10**, 63 (2020).
35. Ambler, R. P. The structure of beta-lactamases. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **289**, 321–331 (1980).
36. Bernabé, K. J., Langendorf, C., Ford, N., Ronat, J.-B. & Murphy, R. A. Antimicrobial resistance in West Africa: a systematic review and meta-analysis. *Int. J. Antimicrob. Agents* **50**, 629–639 (2017).
37. Mahamat, O. O. *et al.* Epidemiology and prevalence of extended-spectrum β-lactamase-and carbapenemase-producing Enterobacteriaceae in humans, animals and the environment in West and Central Africa. *Int. J. Antimicrob. Agents* **57**, 106203 (2021).
38. Hendriksen, R. S. *et al.* Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. *Nat. Commun.* **10**, 1124 (2019).
39. Pärnänen, K. M. M. *et al.* Antibiotic resistance in European wastewater treatment plants mirrors the pattern of clinical antibiotic resistance prevalence. *Sci. Adv.* **5**, eaau9124 (2019).
40. European Centre for Disease Prevention and Control. *Antimicrobial resistance in the EU/EEA (EARS-Net) - Annual Epidemiological Report 2019.* https://www.ecdc.europa.eu/en/publications-data/surveillance-antimicrobial-resistance-europe-2019 (2020).
41. Buelow, E. *et al.* Hospital discharges in urban sanitation systems: Long-term monitoring of wastewater resistome and microbiota in relationship to their eco-exposome. *Water Res. X* **7**, 100045 (2020).
42. Mölder, F. *et al.* Sustainable data analysis with Snakemake. *F1000Research* **10**, 33 (2021).
43. Andrews, S. FastQC A quality control tool for high throughput sequence data https://www.bioinformatics.babraham.ac.uk/projects/fastqc/. (2010).
44. Ewels, P., Ns Magnusson, M., Lundin, S. & Aller, M. K. Data and text mining MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048.
45. Marcel, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).
46. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
47. Bortolaia, V. *et al.* ResFinder 4.0 for predictions of phenotypes from genotypes. *J. Antimicrob. Chemother.* **75**, 3491–3500 (2020).
48. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–9 (2009).
49. Pärnänen, K. *et al.* Maternal gut and breast milk microbiota affect infant gut antibiotic resistome and mobile genetic elements. *Nat. Commun.* **9**, 389 (2018).
50. Beghini, F. *et al.* Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *Elife* **4**, 1–42 (2021).
51. Bengtsson-Palme, J. *et al.* metaxa2: Improved identification and taxonomic classification of small and large subunit rRNA in metagenomic data. *Mol. Ecol. Resour.* **15**, 1403–1414 (2015).
52. Eddy, S. R. Accelerated Profile HMM Searches. *Cit. Eddy SR* **7**, 1002195 (2011).
53. El-Gebali, S. *et al.* The Pfam protein families database in 2019. *Nucleic Acids Res.* **47**, 427–432 (2018).
54. Li, D., Liu, C.-M., Luo, R., Sadakane, K. & Lam, T.-W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**, 1674–1676 (2015).

55. Altschup, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic Local Alignment Search Tool. *J. Mol. Biol* **215**, 403–410 (1990).
56. Murat Eren, A. *et al.* Community-led, integrated, reproducible multi-omics with anvi'o. *Nat. Microbiol.* **6**, 3–6 (2021).
57. Hyatt, D. *et al.* Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, (2010).
58. Wick, R. R., Schultz, M. B., Zobel, J. & Holt, K. E. Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics* **31**, 3350–3352 (2015).
59. Wickham, H. ggplot2: Elegant graphics for data analysis. *Springer-Verlag New York* https://ggplot2.tidyverse.org (2016).
60. Kolde, R. pheatmap: Pretty Heatmaps. R package version 1.0.12. https://cran.r-project.org/package=pheatmap (2019).
61. Pedersen, T. L. patchwork: The Composer of Plots. R package version 1.1.1. https://cran.r-project.org/package=patchwork (2020).
62. South, A. rnaturalearth: World Map Data from Natural Earth. R package version 0.1.0. https://cran.r-project.org/package=rnaturalearth.
63. Venables, W. N. & Ripley, B. Modern Applied Statistics with S. https://www.stats.ox.ac.uk/pub/MASS4 (2002).
64. McMurdie, P. J. & Holmes, S. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS One* **8**, (2013).
65. Oksanen, J. *et al.* vegan: Community Ecology Package. https://cran.r-project.org/package=vegan (2020).
66. Chen, H. VennDiagram: Generate High-Resolution Venn and Euler Plots. R package version 1.6.20. https://cran.r-project.org/package=VennDiagram (2018).
67. Revelle, W. psych: Procedures for Psychological, Psychometric, and Personality Research. https://cran.r-project.org/package=psych (2021).
68. Kassambara, A. ggcorrplot: Visualization of a Correlation Matrix using 'ggplot2'. https://cran.r-project.org/package=ggcorrplot (2019).
69. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
70. Wisplinghoff, H. *et al.* Related Clones Containing SCCmec Type IV Predominate among Clinically Significant Staphylococcus epidermidis Isolates. *Antimicrob. Agents Chemother.* **47**, 3574–3579 (2003).
71. Naas, T. *et al.* Beta-lactamase database (BLDB)-structure and function. *J. Enzyme Inhib. Med. Chem.* **32**, 917–919 (2017).
72. Fu, L., Niu, B., Zhu, Z., Wu, S. & Li, W. Sequence analysis CD-HIT: accelerated for clustering the next-generation sequencing data. **28**, 3150–3152 (2012).
73. Thompson, L. *et al.* A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* **551**, 457–463 (2017).
74. Oyem, I. M., Oyem, H. H. & Atuanya. Molecular characterisation of bacteria strains of septic tank sewage samples from related sites in Delta and Edo States of Nigeria using 16S rRNA denaturing gradient gel electrophoresis (DGGE). **14**, 132–138 (2020).
75. Laursen, M. F. *et al.* Faecalibacterium Gut Colonization Is Accelerated by Presence of Older Siblings. *mSphere* **2**, e00448-17 (2017).
76. Périchon, B. *et al.* Identification of 50 class D β-lactamases and 65 Acinetobacter-derived cephalosporinases in Acinetobacter spp. *Antimicrob. Agents Chemother.* **58**, 936–949 (2014).
77. Figueiredo, S., Bonnin, R. A., Poirel, L., Duranteau, J. & Nordmann, P. Identification of the naturally occurring genes encoding carbapenem-hydrolysing oxacillinases from *Acinetobacter haemolyticus*, *Acinetobacter johnsonii*, and *Acinetobacter calcoaceticus*. *Clin Microbiol Infect* **18**, 907–913 (2012).
78. Cox, G. & Wright, G. D. Intrinsic antibiotic resistance: Mechanisms, origins, challenges and solutions. *Int. J. Med. Microbiol.* **303**, 287–292 (2013).
79. Araújo, S. *et al.* Carbapenem-resistant bacteria over a wastewater treatment process: Carbapenem-resistant *Enterobacteriaceae* in untreated wastewater and intrinsically-resistant bacteria in final effluent. *Sci. Total Environ.* **782**, 146892 (2021).
80. Zieliński, W. *et al.* Wastewater treatment plants as a reservoir of integrase and antibiotic resistance genes – An epidemiological threat to workers and environment. *Environ. Int.* **156**, 106641 (2021).
81. Pal, C., Bengtsson-Palme, J., Kristiansson, E. & Larsson, D. G. J. Co-occurrence of resistance genes to antibiotics, biocides and metals reveals novel insights into their co-selection potential. *BMC Genomics* **16**, (2015).
82. Gillings, M. R., Xuejun, D., Hardwick, S. A., Holley, M. P. & Stokes, H. W. Gene cassettes encoding resistance to quaternary ammonium compounds: a role in the origin of clinical class 1 integrons? *ISME*

J. **3**, 209–215 (2009).
83. Olalekan, A. *et al.* High proportion of carbapenemase-producing *Escherichia coli* and *Klebsiella pneumoniae* among extended-spectrum β-lactamase-producers in Nigerian hospitals. *J. Glob. Antimicrob. Resist.* **21**, 8–12 (2020).
84. Manenzhe, R. I., Zar, H. J., Nicol, M. P. & Kaba, M. The spread of carbapenemase-producing bacteria in Africa: a systematic review. *J. Antimicrob. Chemother.* **70**, 23–40 (2015).
85. van Duin, D. & Doi, Y. The global epidemiology of carbapenemase-producing Enterobacteriaceae. *Virulence* **8**, 460–469 (2017).
86. Mushi, M. F., Mshana, S. E., Imirzalioglu, C. & Bwanga, F. Carbapenemase Genes among Multidrug Resistant Gram Negative Clinical Isolates from a Tertiary Hospital in Mwanza, Tanzania. *Biomed Res. Int.* **2014**, 303104 (2014).
87. Yehouenou, C. *et al.* First detection of a plasmid-encoded New-Delhi metallo-beta-lactamase-1 (NDM-1) producing *Acinetobacter baumannii* using whole genome sequencing, isolated in a clinical setting in Benin. *Ann. Clin. Microbiol. Antimicrob.* **20**, 5 (2021).
88. Wangai, F. K., Masika, M. M., Maritim, M. C. & Seaton, R. A. Methicillin-resistant *Staphylococcus aureus* (MRSA) in East Africa: red alert or red herring? *BMC Infect. Dis.* **19**, 596 (2019).
89. Garoy, Y. E. *et al.* Methicillin-Resistant Staphylococcus aureus (MRSA): Prevalence and Antimicrobial Sensitivity Pattern among Patients-A Multicenter Study in Asmara, Eritrea. **6**, 8321–834 (2019).
90. Falagas, M. E., Karageorgopoulos, D. E., Leptidis, J. & Korbila, I. P. MRSA in Africa: Filling the Global Map of Antimicrobial Resistance. *PLoS One* **8**, 68024 (2013).
91. The Center for Disease Dynamics Economics & Policy. ResistanceMap 2018. *The Center for Disease Dynamics Economics & Policy. ResistanceMap.* https://resistancemap.cddep.org/AntibioticResistance.php.
92. Mutai, W. C. *et al.* High Prevalence of Multidrug-Resistant *Clostridioides difficile* Following Extensive Use of Antimicrobials in Hospitalized Patients in Kenya. *Front. Cell. Infect. Microbiol.* **10**, 897 (2021).
93. Van, T. T. H., Yidana, Z., Smooker, P. M. & Coloe, P. J. Antibiotic use in food animals worldwide, with a focus on Africa: Pluses and minuses. *J. Glob. Antimicrob. Resist.* **20**, 170–177 (2020).
94. Olumuyiwa Olaitan, A. *et al.* Banning colistin in feed additives: a small step in the right direction. *Lancet Infect. Dis.* **21**, 29–30 (2021).
95. Olowo-Okere, A. & Yacouba, A. Review Molecular mechanisms of colistin resistance in Africa: A systematic review of literature. *Germs* **10**, 367–379 (2020).
96. Kneis, D., Berendonk, T. U. & Heß, S. High prevalence of colistin resistance genes in German municipal wastewater. *Sci. Total Environ.* **694**, 133454 (2019).
97. Ngbede, E. O. *et al.* Identification of mobile colistin resistance genes (*mcr-1.1, mcr-5 and mcr-8.1*) in *Enterobacteriaceae* and *Alcaligenes faecalis* of human and animal origin, Nigeria. *Int. J. Antimicrob. Agents* **56**, 106108 (2020).
98. Borowiak, M. *et al.* Identification of a novel transposon-associated phosphoethanolamine transferase gene, *mcr-5*, conferring colistin resistance in d-tartrate fermenting *Salmonella enterica* subsp. *enterica* serovar Paratyphi B. *J. Antimicrob. Chemother.* **72**, 3317–3324 (2017).
99. Borowiak, M. *et al.* Characterization of *mcr-5*-Harboring *Salmonella enterica* subsp. *enterica* Serovar Typhimurium Isolates from Animal and Food Origin in Germany. *Antimicrob. Agents Chemother.* **63**, e00063-19 (2019).
100. Guo, S. *et al.* Conjugative IncX1 plasmid harboring colistin resistance gene *mcr-5.1* in *Escherichia coli* isolated from chicken rice retailed in Singapore. *Antimicrob. Agents Chemother.* **63**, e01043-19 (2019).
101. Peace Hounkpe, S., Codjo Adjovi, E., Crapper, M. & Awuah, E. Wastewater Management in Third World Cities: Case Study of Cotonou. *Benin. J. Environ. Prot.* **5**, 387–399 (2014).
102. Bougnom, B. P. *et al.* Wastewater used for urban agriculture in West Africa as a reservoir for antibacterial resistance dissemination. *Environ. Res.* **168**, 14–24 (2019).
103. Bougnom, B. P., Thiele-Bruhn, S., Ricci, V., Zongo, C. & Piddock, L. J. V. Raw wastewater irrigation for urban agriculture in three African cities increases the abundance of transferable antibiotic resistance genes in soil, including those encoding extended spectrum β-lactamases (ESBLs). *Sci. Total Environ.* **698**, 134201 (2020).
104. Totin, H. S. *et al.* Groundwater pollution and the safe water supply challenge in Cotonou town, Benin (West Africa). *IAHS-AISH Proc. Reports* **361**, 191–196 (2013).

For Table of Contents Only