

# Journal of Advanced Pharmacy Research

## Section E: Microbiology & Immunology



### Superbug Arsenal: Antimicrobial Resistance in Hospital and Wastewater Environments

Ahmed I. Ahmed<sup>1</sup>, Ahmed H. Saleh<sup>1</sup>, Michael M. Kamal<sup>1</sup>, Mennatallah M. Abdelrahman<sup>1</sup>, Salma Abou Elhassan<sup>1</sup>, Rana S. Khalel<sup>1</sup>, Ali H. A. Elbehery<sup>2\*</sup>

<sup>1</sup>Department of Biology, The American University in Cairo, 11835 Cairo, Egypt.

<sup>2</sup>Department of Microbiology and Immunology, Faculty of Pharmacy, University of Sadat City, 32897 Sadat City, Egypt.

\*Corresponding author: Ali H.A. Elbehery, Department of Microbiology and Immunology, Faculty of Pharmacy, University of Sadat City, 32897 Sadat City, Egypt. Tel. (+2)01008714597  
Email address: [ali.elbehery@fop.usc.edu.eg](mailto:ali.elbehery@fop.usc.edu.eg)

Submitted on: 17-03-2023; Revised on: 19-04-2023; Accepted on: 22-04-2023

**To cite this article:** Ahmed, A. I.; Saleh, A. H.; Kamal, M. M.; Abdelrahman, M. M.; Abou Elhassan, S.; Khalel, R. S.; Elbehery, A. H. Superbug Arsenal: Antimicrobial Resistance in Hospital and Wastewater Environments. *J. Adv. Pharm. Res.* **2023**, 7 (3), 170-180. DOI: [10.21608/aprh.2023.200708.1215](https://doi.org/10.21608/aprh.2023.200708.1215)

#### ABSTRACT

**Background:** Antimicrobial resistance (AMR) is a serious problem with growing threats to the health situation of societies. AMR is particularly important in hospital environments and is a major cause of death in nosocomial (hospital-acquired) infections. Similarly, wastewater is a hotspot for AMR with high bacterial diversity and stress exerted by pollutants, such as antibiotics and heavy metals. **Objectives:** We sought to get a better insight of AMR in these two environments. **Methods:** In 30 high-throughput sequenced metagenomes from both environments, we assessed AMR using the open reading frames, called from the assembled contigs, via Resistance Gene Identifier. Taxonomy of the sequences for which AMR could be detected was also determined. **Results:** Taxonomy generally agreed between both environments for the most abundant families, namely, Enterobacteriaceae, Moraxellaceae, and Pseudomonadaceae. However, for other less abundant families, each environment showed unique distribution. On the genus level, Moraxellaceae and Pseudomonadaceae were mainly dominated by *Acinetobacter* and *Pseudomonas*, respectively, which are notorious superbugs. Assessment of AMR identified 1,652 AMR genes belonging to 156 gene families. Tetracycline occupied the top of the list of drug classes for which resistance was detected in both environments, followed by fluoroquinolones in hospital samples and macrolides in wastewater samples. Top resistance mechanisms detected in both environments were efflux pumps and antibiotic inactivation. Accordingly, gene families manifested this pattern of resistance mechanisms with the top three gene families being Resistance-Nodulation-Division (RND) efflux pumps, Major Facilitator Superfamily (MFS) efflux pumps, and OXA-beta-lactamases. **Conclusions:** Taken together, we shed light on AMR in two particularly important environments, emphasizing the significance of wastewater as a reservoir of resistance.

**Keywords:** Antimicrobial resistance, Hospitals, Wastewater, Metagenomics, High-throughput Sequencing, Superbugs

#### INTRODUCTION

Antimicrobial resistance has emerged as one of the most serious public health issues of the twenty-first

century due to several reasons, including increased antibiotic consumption. In the last 20 years, antibiotic consumption has risen by 65% and is depicted to increase by 200% by 2030<sup>1</sup>. As resistance develops and

spreads, the treatment of several common infections is rendered extremely difficult. Recently, the number of deaths caused by resistant infections has witnessed a dramatic increase to 1.27 million deaths per year <sup>2</sup>. Today, there is an emerging health threat through superbugs, which are extremely resistant bacteria that show resistance against multiple classes of antibiotics. The most notable superbugs are *Enterococcus* spp., *Staphylococcus aureus*, *Klebsiella pneumoniae* (*K. pneumoniae*), *Acinetobacter baumannii* (*A. baumannii*), *Pseudomonas aeruginosa*, and *Enterobacter* spp., abbreviated as ESKAPE, which are a group of life-threatening pathogens that are the leading causes of nosocomial infections worldwide <sup>3</sup>.

Hospital environments are hot spots for antimicrobial resistance owing to high levels of antimicrobial consumption <sup>4</sup>. The most common species of antibiotic-resistant organisms causing infections in hospitals are methicillin-resistant *Staphylococcus aureus* (MRSA), non-typhoidal *Salmonella* (NTS), *K. pneumoniae*, and *Acinetobacter baumannii*. Invasive medical devices and surgical procedures are associated with many of these hospital infections because they allow free access of bacteria to the patient's body. For instance, *A. baumannii* is responsible for up to 20% of infections in intensive care units since it attaches to medical devices and causes devastating infections in lungs, urinary tracts, and bloodstream, especially among susceptible and immunocompromised patients <sup>5</sup>.

Similarly, the aggregation of wastewater enriched with metals, antibiotic residues, bacteria, and nutrients at wastewater treatment plants (WWTPs) makes it a hotspot for antibiotic resistance that facilitates the propagation of antibiotic-resistance genes (ARGs) <sup>4</sup>. Micropollutants such as chemicals, pharmaceuticals, ingredients of households, environmental persistent pollutants substances (EPPS), or pesticides may not be abolished in conventional water treatment processes, leading to water pollution. Although these substances will have quite low concentrations; yet these substances and mostly EPPSs could act as endocrine-disrupting substances, genotoxic substances, and enhancers of bacterial resistances. Inadequate treatment of wastewater effluents and its irresponsible disposal leads to the modulation of bacterial genomes expression that is responsible for the increase of the AMR phenomenon <sup>6</sup>.

Herein, we aim to identify the antibiotic-resistant genes across a wide variety of selected metagenomic samples obtained from various ecological niches that are pertaining to hospital and WWTPs environments. The assembly of the resulting contigs from such data were analyzed using Resistance Gene Identifier (RGI) algorithm that is a part of the Comprehensive Antibiotic Resistance Database (CARD).

## MATERIAL AND METHODS

### Obtaining the sequences of the samples

The raw sequence reads of the chosen metagenomic samples obtained from hospital environment and wastewater treatment plant environment were downloaded from NCBI Sequence Read Archive (SRA) <sup>7</sup> using the prefetch tool v2.3.5 of the SRA toolkit. All the chosen samples were paired-end sequences and had an average size of one GB per metagenome. FASTQ files were obtained using the fasterq-dump tool of the SRA toolkit. The metagenomic samples for the hospital environment were obtained from seven studies that had their metagenomic samples collected from microbial communities in the hospital environment with a total of 17 samples selected. The metagenomic samples for the wastewater treatment plant environment were obtained from four studies that had their metagenomic samples collected from microbial communities in the wastewater treatment plant environment with a total of 13 samples selected. In total, the study addresses 30 metagenomic samples from both environments. **Table S1** outlines the studies selected for the hospital and wastewater environments.

### Quality control

Quality control was done using fastp tool v0.20.0 <sup>8</sup> to improve the overall quality of the sequences by removing low quality sequence reads, trimming low quality ends, and adapters, and deduplication to remove redundant sequences. Parameters were set to defaults. Default parameters allow no more than 40% of bases per read to fall below a threshold quality phred score of 15. The quality of the generated reads was evaluated using FastQC v0.11.3 <sup>9</sup> to assess the length of the sequences, the quality, the GC content, the N content, and the duplication level.

### Assembly

Assembly of the short sequence reads was done using MEGAHIT v1.2.9 <sup>10</sup>. Parameters were set at defaults with kmer size ranging between 21 and 141.

### Taxonomic profiling

Taxonomic profiling of the sequences was done using Kraken v1.1.1 <sup>11</sup>. Kraken is a kmer-based alignment tool, used to identify taxonomy in shotgun metagenomic sequences. Two Kraken commands were also used, namely kraken-translate and kraken-report. The translate command was used to add the full taxonomy to the output file since the default output contains the taxonomy ID only. The kraken-report command was also used to make a summarized report of all the assigned taxonomies across each sample.

## AMR detection

The detection of antimicrobial resistance genes was done using CARD Resistance Gene Identifier (RGI) <sup>12</sup>. RGI is a widely used tool for the identification of resistance genes and has been used for metagenomes in several studies <sup>13-15</sup>. The --alignment\_tool option was used to specify the alignment tool, which was specified as Diamond v0.9.30 <sup>16</sup>. Diamond was used to align the protein sequences predicted from the sample against the CARD database. The heatmap option was also used with RGI to allow multiple analysis and visualization of the results. Two options were used with rgi heatmap which are --category and -clus. The --category option was used to organize the visualization of the identified resistant genes in the heatmap based on a specific category and the category chosen is resistance mechanism. The -clus option was used to cluster the resistance genes of each sample together improving the visualization quality of the heatmap. Moreover, Seqtk <sup>17</sup> tool was used to retrieve the nucleotide sequences of the obtained anti-microbial resistance genes from the original contigs file. BBmap <sup>18</sup> was used to map the quality-controlled sequence reads of the samples to the contigs of the detected AMR genes to assess the abundance of these genes. Two options were used with BBmap scaffstats and covstats, which were used to obtain statistics about the mapping of the reads to the contigs and their coverage. Lastly, the function and the taxonomic classification of the detected AMR genes were retrieved from the comprehensive annotation and taxonomic classification files that were generated earlier using SEED and kraken. respectively.

## RESULTS

### Overview of the metagenomic samples

#### Sequence reads

Metagenomic samples were obtained from SRA. Overall, they were 30 samples with a total number of reads of 618,772,582 and a total number of bases of 89.6 Gbp, before quality control (**Table S2**). More than 590 million reads passed quality control with a total number of bases of 86.08 Gbp (96%). The highest (41,903,544) and lowest (5,037,848) number of high-quality reads in the hospital environment were found in the samples SRR11745697 and SRR13893607, respectively. While in the wastewater environment, the highest (37,538,086) and lowest (16,904,628) number of high-quality reads in the hospital environment were found in the samples SRR14932565 and SRR16214402, respectively.

#### Contigs

The assembly of high quality gave rise to contigs of variable lengths (**Table S3**). The highest number of contigs obtained in the hospital environment

was 843,697 contigs, found in the sample SRR11745693, yet the average contig length was the lowest in this environment (455.69 bp and N50 = 450 bp). The lowest number of contigs in the hospital environment was 70 contigs in the sample SRR13893626. However, the average contig length was the highest in this environment (71,400.13 bp and N50 = 369,951 bp). The wastewater environment showed a similar pattern with the highest number of contigs obtained was in the sample SRR14932610 being 559,379 contigs, yet the average contig length (441 bp and N50 = 665 bp) was among the lowest average contig lengths in this environment. The lowest number of contigs in the wastewater environment was obtained in the sample SRR16214400 and was 1,341 contigs with an average contig length of 1,554 bp and N50 of 137,588 bp, the highest contig length and N50 in this environment.

### Taxonomic Assignment of the hospital and sewage environments

Taxonomic profiling for each environment was essential to determine the abundance of different taxonomic groups within the samples. To be more focused on the antimicrobial resistance impact, the taxonomic profiling was done for the contigs in which AMR genes were detected. **Figure 1** shows the relative abundance of 41 microbial families. The family level was selected because the genus for many of the hits could not be identified. Based on the family level, the percentage of the hits that were classified in the hospital and sewage environments were 76.03% and 57.60% respectively. The relative abundance on the family level differed between the hospital and sewage environments (**Figure 1**). Yet, the three most abundant families in both environments were the same, namely Enterobacteriaceae, Moraxellaceae and Pseudomonadaceae. These three families collectively accounted for 58.54% and 62% of the hospital and sewage samples respectively. The abundance of these three families for the two environments is separately shown in **Figure 2**. Despite the similarities between these three families, there were also some differences in the taxonomic classifications detected in the two environments. For example, the Staphylococcaceae and the Xanthomonadaceae families each contributed to around 6% of the hospital environment samples, however, they were not detected at all in the sewage environment. On the other hand, the sewage environment showed considerably higher abundance of the Aeromonadaceae and Peptostreptococcaceae families, with approximately 8 and 3%, respectively for the sewage environment versus 2 and 0%, respectively for the hospital environment.

**Figure 2** shows that the most abundant family in the case of the hospital environment was the

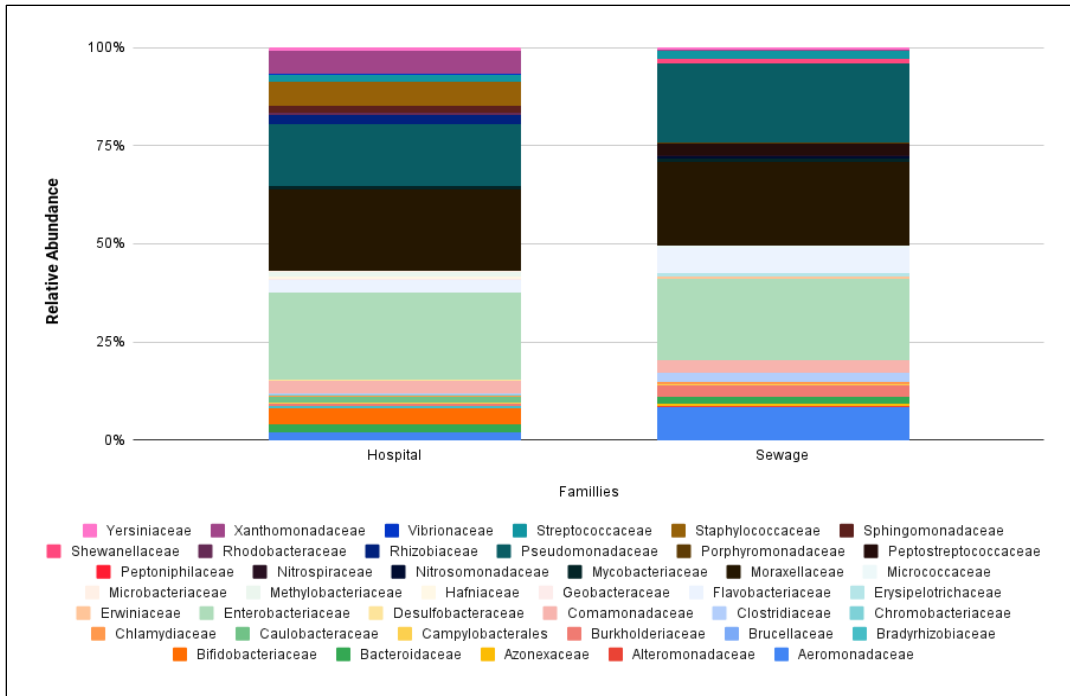


Figure 1. Relative abundance of all the detected microbial taxa, at the family level, of the processed samples per each of the two environments.

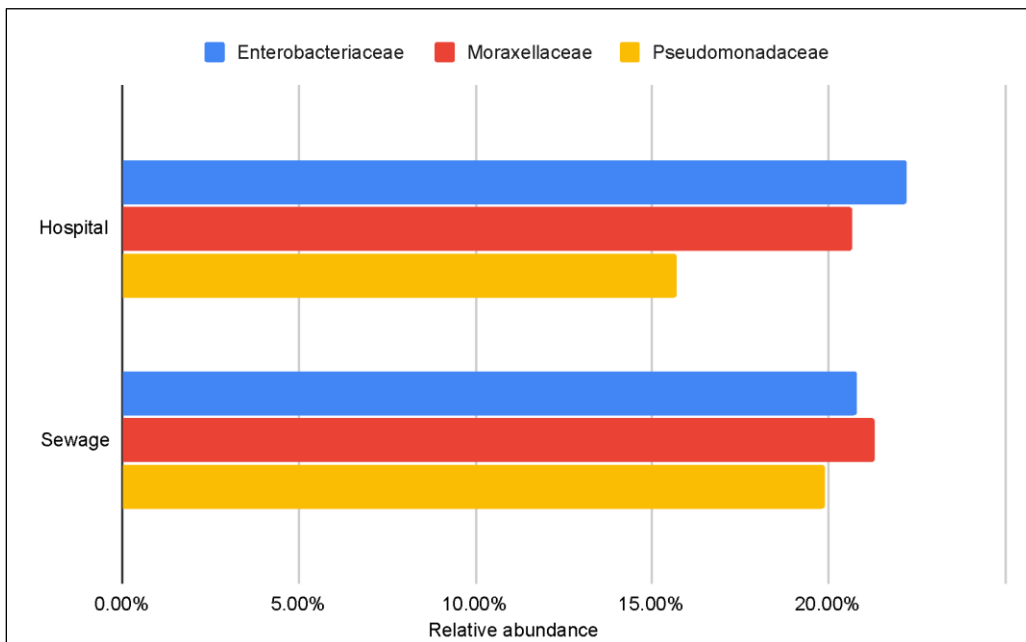


Figure 2. Comparison between the relative abundance of the three selected families per each environment.

Enterobacteriaceae with 22.2%, while the Moraxellaceae family was the most abundant in the sewage environment with 21.3%. However, the abundance of the three families was generally close to each other in both environments. Thus, inspection of the taxonomic classification on the genus level for the candidates of these three families was done to explore whether the two environments had the same level of similarity on that taxonomic level.

Interestingly, for the Pseudomonadaceae family, it was found that in both environments, 100% of the family hits belonged to the *Pseudomonas* genus. On the species level, the most abundant species in both environments was found to be *Pseudomonas aeruginosa*, representing 50 and 56% in the hospital and wastewater environments, respectively out of the identified *Pseudomonas* species in each environment. As for the Moraxellaceae family, all the hits belonged to the *Acinetobacter* genus in both environments except for only one hit in the hospital samples that belonged to the *Psychrobacter* genus. Additionally, the most common species in the case of the hospital environment was the *Acinetobacter baumannii* (68%), while the most common species in the sewage environment was the *Acinetobacter johnsonii* (58%). Finally, the Enterobacteriaceae family showed the highest diversity in terms of identified genera/species. The most common in the hospital environment were the *Escherichia coli* (47%), followed by *Klebsiella* (20%), half of which was identified to be *K. pneumoniae*, followed by *Citrobacter* (15%), most of which (83%) was identified as *Citrobacter freundii*. Similarly, the most common genera/species in the sewage environment were *Escherichia coli* (39%), *Klebsiella* (18%), particularly *K. pneumoniae* (42% of identified *Klebsiella* species), *Citrobacter* (10%), and *Lelliottia* (10%). Some genera e.g., *Lelliottia* and *Raoultella* could be found only in the wastewater samples, but not in hospital samples.

### AMR genes profile

The AMR genes profiles of the 30 metagenomic samples were detected using CARD Resistance Gene Identifier (RGI). Each sample was analyzed according to the AMR gene families they belong to, the drug classes they showed resistance to, and their resistance mechanisms of action. A total number of 1650 AMR genes were detected that belonged to a total number of 102 AMR gene families, conferring resistance to 64 drug classes, and carry out resistance through 13 different resistance mechanisms. In the hospital environment, 17 metagenomic samples were processed. CARD RGI tool detected a total of 1040 AMR genes from 90 different AMR families. These AMR genes confer resistance to a total of 31 different drug classes (Figure 3). Detected resistance in the hospital environment belonged to six different

mechanisms (Figure 4). There are six major drug classes to which detected AMR genes in the hospital environment confer resistance. These six classes represent 55.95% of all drug classes to which AMR genes in the hospital environment confer resistance. These classes, in a descending order of abundance, are tetracyclines (11.78%), fluoroquinolones (9.87%), cephalosporins (9.39%), penams (9.05%), aminoglycosides (8.34%), and macrolides (7.52%) (Figure 3). The most abundant mechanisms of action in the hospital environment were antibiotic inactivation (38.94%), followed by antibiotic efflux (35.58%), while the least abundant were reduced permeability to antibiotics (1.18%) and antibiotic target protection (6.82%) (Figure 4). On the other hand, 610 AMR genes could be detected in the processed samples of the wastewater environment. These AMR genes can be categorized into a total of 65 AMR families. They confer resistance to a total of 30 different drug classes (Figure 3). Similar to the hospital environment, the detected AMR genes in the wastewater environment belonged to six different mechanisms of resistance (Figure 4). The top six drug classes in terms of abundance were as follows: tetracyclines (13.39%), macrolides (10.42%), fluoroquinolones (9.21%), penams (8.42%), cephalosporins (8.24%), and phenicol antibiotics (6.18%) (Figure 3). These six classes collectively accounted for 55.86% of the total abundance of resistance detected in the wastewater environment. Concerning the mechanisms of resistance, the most abundant ones in the wastewater environment were antibiotic efflux (40.26%), followed by antibiotic inactivation (28.65%), while the least abundant were reduced permeability to antibiotics (1.15%) and antibiotic target replacement (3.44%).

Concerning the detected AMR gene families, the top five in terms of abundance had a similar order in both environments (Figure 5, Table S4). The most abundant AMR gene family was the resistance-nodulation-cell division (RND) antibiotic efflux pump with abundances of 21.26% and 23.97% in the hospital and wastewater environments, respectively. RND efflux pump was followed by major facilitator superfamily (MFS) antibiotic efflux pump with abundances of 12.39% and 14.29% in the hospital and wastewater environments, respectively. The only difference in order between both environments was between aminoglycoside nucleotidyltransferase ANT(3'') and tetracycline-resistant ribosomal protection protein. ANT(3'') was higher in abundance than tetracycline-resistant ribosomal protection protein in the hospital environment (5.16% vs 3.18%, respectively), while slightly lower than tetracycline-resistant ribosomal protection protein in the wastewater environment (3.78% vs 4.11%, respectively).



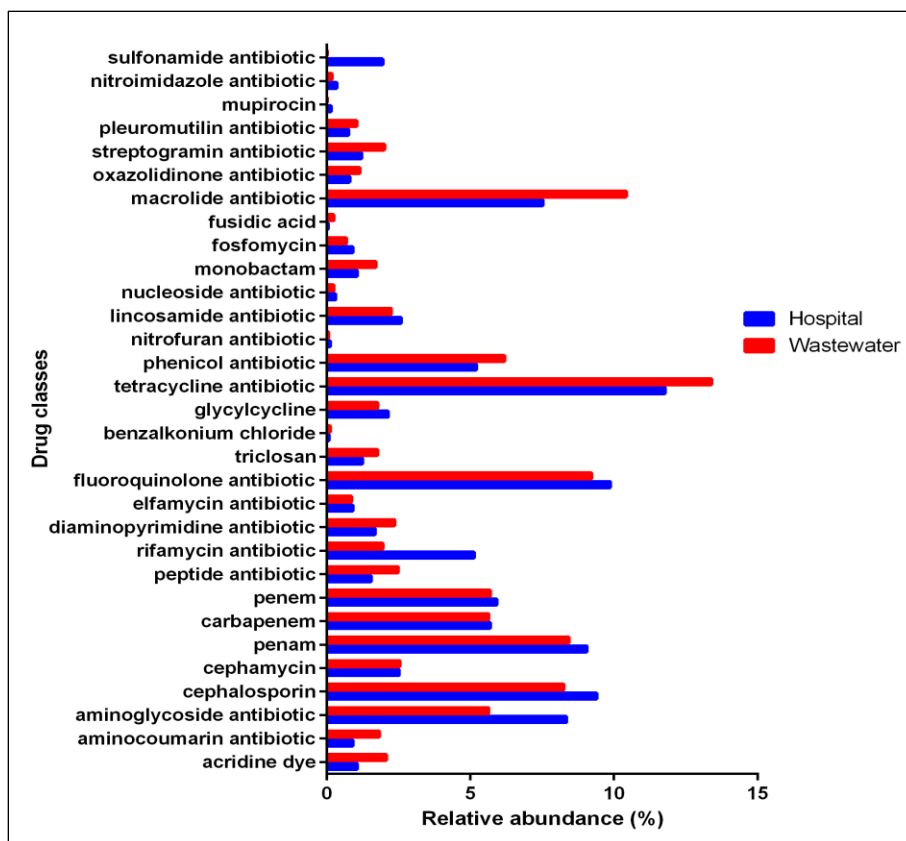


Figure 3. Relative abundance of resistance genes according to drug classes they confer resistance to.

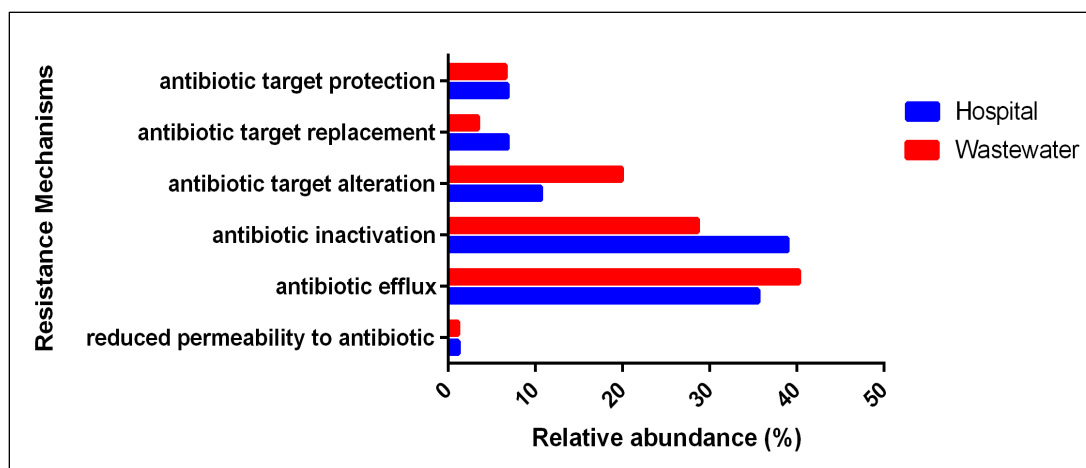


Figure 4. Relative abundance of resistance genes according to resistance mechanisms

Principal component analysis (PCA) for resistance profiles in both environments based on the relative abundance of resistance genes for the different drug classes for which resistance was detected (Figure 6) showed that a total of 50.11% of variance can be explained by principal components (PCs) 1 and 2; PC1,

36.43% and PC2, 13.67%. It can be observed that samples from each environment roughly cluster together on the PCA plot, which probably denotes that despite the similarity between both environments for the most abundant classes, the overall profile is distinct for each environment.

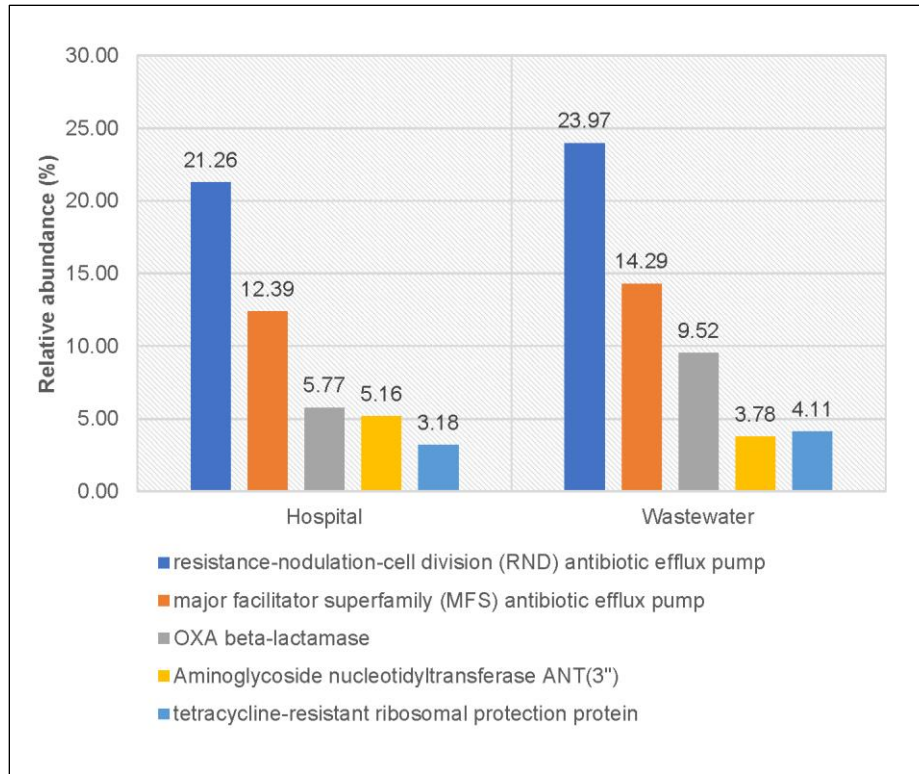


Figure 5. Clustered bar chart showing the abundance of the top five detected gene families in the hospital and wastewater environments.

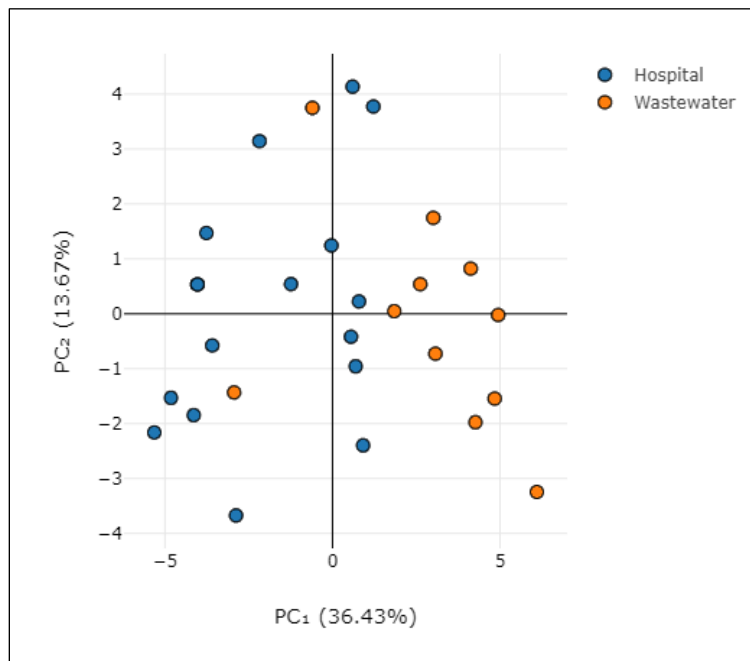


Figure 6. Principal component analysis (PCA) for resistance profiles in both environments based on the relative abundance of resistance genes for the different drug classes for which resistance was detected.

## DISCUSSION

In this study, we obtained 30 different metagenomic samples from hospitals and wastewater environments. Both environments are hotspots for AMR<sup>4</sup>; therefore, investigation of different samples is warranted and would potentially enhance our understanding of resistance in these environments. Analysis of the samples included not only the detection of AMR genes, but also the taxonomic assignment thereof. Taxonomic assignment revealed that the three most abundant bacterial families in both environments were Enterobacteriaceae, Moraxellaceae and Pseudomonadaceae. Interestingly, Enterobacteriaceae as well as Pseudomonadaceae were previously reported to be among the most frequently detected families in the hospital environment<sup>19, 20</sup>. Similarly, Enterobacteriaceae, Moraxellaceae and Pseudomonadaceae were among the most frequently isolated resistant bacterial families detected in lentic and effluent water (wastewater) in Delhi National Capital Region, India<sup>21</sup>. Another study also showed that Enterobacteriaceae and Pseudomonadaceae were the most abundant resistant families isolated from wastewater in two different WWTPs in Charlotte, NC, USA<sup>22</sup>. On the genus/species level, particular genera and/or species could be detected, including *Escherichia coli*, *Klebsiella pneumoniae* (Enterobacteriaceae), *Pseudomonas aeruginosa* (Pseudomonadaceae), and *Acinetobacter* (Moraxellaceae). These bacteria are typically found amongst the most frequently isolated resistant bacteria in hospitals and clinical settings<sup>23-25</sup>. Similarly, all these genera, which are among the World Health Organization (WHO) priority list of resistant bacteria, are constantly being reported among the resistant bacteria detected in wastewater environments<sup>26, 27</sup>. Both *Lelliottia* and *Raoultella*, which were detected only in wastewater samples in this study, are Gram-negative bacteria belonging to the Enterobacteriaceae family. They are usually found in natural environments, such as water and soil. They are also opportunistic human pathogens. For both, antimicrobial resistance phenotypes, particularly beta-lactamase resistance, as well as multidrug resistance have been reported<sup>28, 29</sup>.

Concerning resistance gene profiles for the studied samples, a total of 1650 resistance genes were detected. These genes were highly diverse in terms of the drug classes they confer resistance to, and the resistance mechanisms and the gene families they belong to. In the hospital environment, 1040 AMR genes could be detected. These genes confer resistance to 31 different drug classes. The top six drug classes in terms of their corresponding AMR gene abundance were tetracyclines, fluoroquinolones, cephalosporins, penams, aminoglycosides, and macrolides. According

to the WHO, all these classes are categorized as critically important antimicrobials, except tetracyclines, which are considered highly important antimicrobials. This categorization is based on the fulfillment of two criteria: C1) the antimicrobial class is the only available therapy, or one of the few available to treat severe human bacterial infections; C2) the antimicrobial class is used to treat bacterial infections that can either be transmitted to humans from non-human sources or acquire resistance genes from non-human sources. Antimicrobial classes that fulfill these two criteria are referred to as critically important, while those which fulfill only one of the two criteria are referred to as highly important. It is evident that both categories (critically and highly important) are highly amenable to resistance and should be prudently used in clinical practice<sup>30</sup>. Similarly, in the wastewater samples, 610 AMR genes could be detected, conferring resistance to 30 different drug classes. The top of the list of these drug classes included tetracyclines, macrolides, fluoroquinolones, penams, cephalosporins, and phenicol antibiotics. According to the WHO list, only tetracyclines and phenicol antibiotics are considered highly important, while the rest are regarded as critically important<sup>30</sup>. Tetracyclines, the most abundant type of resistance in both hospital and wastewater samples in this study, find widespread use clinically and in agriculture. This extensive use maintains a selective pressure, contributing to the development of resistance<sup>31</sup>. Therefore, several studies report tetracyclines as the most abundant type of resistance. For example, Maestre-Carballa, et al.<sup>32</sup> have recently reported tetracyclines to be, on average, the most abundant type of resistance in the Human Microbiome Project samples. Similarly, in wastewater tetracyclines were also found to be most abundant in sludge samples in sewage treatment plants<sup>33, 34</sup>.

The most abundant mechanisms of resistance in both of the studied environments were antibiotic efflux and antibiotic inactivation. Consistent with these results, a recent study, conducted for all complete RefSeq bacterial genomes (15,790), showed that the most abundant mechanism of resistance, based on the number of detected resistant loci, was antibiotic efflux, followed by antibiotic inactivation<sup>35</sup>. Many of these pumps are natively encoded by immobilized, chromosomal genes, imparting intrinsic resistance to different classes of antibiotics, as well as other chemicals. Unlike efflux, antibiotic inactivation was found to cause resistance for numerous classes of antibiotics with genes being mobilized via plasmids and insertion sequences<sup>35</sup>. The top five gene families in both environments reflected the abundance of the antibiotic efflux and antibiotic inactivation mechanisms. The top two were RND and MFS efflux pumps, while the rest were different antibiotic-inactivating enzymes,



namely OXA beta-lactamase, aminoglycoside nucleotidyltransferase ANT(3") and tetracycline-resistant ribosomal protection protein, catalyzing the inactivation of aminoglycosides, beta lactam antibiotics and tetracyclines, respectively. It is worth noting that clinically significant antibiotic efflux resistance in Gram-negative and Gram-positive bacteria is usually mediated by the RND and MFS efflux pumps, respectively <sup>36</sup>, a note that highlights the significance and the role of the gene families of the highest abundance detected in the studied samples in potential infections.

Generally, wastewater samples showed a similar profile to those of the hospital samples, especially with regards to the most abundant drug classes for which resistance was shown. This could be explained by the presence of a similar anthropogenic impact in both environments. Similar to the antibiotic stress found in hospitals, wastewater also exhibits such stress since significant portions of the antibiotics given to humans or animals are excreted into the sewage because they are not metabolized. Therefore, it is not uncommon to find the ARGs detected in wastewater, contained in clinically relevant pathogenic bacteria <sup>37</sup>. Recently, wastewater metagenomic data have been used to predict the prevalence of clinical resistance and the resistance data inferred from the wastewater metagenomes correlated well with clinical surveillance data <sup>38</sup>.

## CONCLUSION

Overall, this study highlighted the presence of several opportunistic and pathogenic bacteria in hospital and wastewater environments. These bacteria harbored diverse ARGs that confer resistance to different classes of antimicrobials, use different mechanisms and belong to different gene families. Nevertheless, both environments showed a rather similar profile that emphasizes the importance of wastewater resistance as a predictor and a potential reservoir of resistance in clinical settings.

## Funding Acknowledgment

No external funding was received.

## Conflict of interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

## Acknowledgement

The authors would like to deeply thank Prof. Ahmed Moustafa, Professor of Bioinformatics and Chair of the Department of Biology, The American University in Cairo, for his help in study coordination and provision of the required computational resources.

## Supplementary Material

**Table S1.** Selected metagenomic projects for the hospital and wastewater environments.

**Table S2.** Statistics for studied sample sequence reads before and after quality control.

**Table S3.** Statistics for generated assembled contigs.

**Table S4.** Relative abundance of detected AR gene families in the hospital and wastewater environments.

**List of Abbreviations:** AMR, Antimicrobial resistance; ANT, Aminoglycoside nucleotidyltransferase; ARGs, Antibiotic resistance genes; CARD, Comprehensive Antibiotic Resistance Database; EPPS, Environmental persistent pollutants; MFS, Major Facilitator Superfamily; MRSA, Methicillin-resistant *Staphylococcus aureus*; NCBI, National Center for Biotechnology Information; NTS, Non-typhoidal *Salmonella*; PCA, Principal component analysis; PC, Principal component; RGI, Resistance Gene Identifier; RND, Resistance-Nodulation-Division; SRA, Sequence Read Archive; WHO, World Health Organization; WWTPs, Wastewater treatment plants.

## REFERENCES

1. Dadgostar P. Antimicrobial Resistance: Implications and Costs. *Infect Drug Resist.* **2019**, *12*, 3903-3910, <https://doi.org/10.2147/idr.S234610>.
2. Murray CJL; Ikuta KS; Sharara F; Swetschinski L; Robles Aguilar G; Gray A, et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet.* **2022**, *399*(10325), 629-655, [https://doi.org/10.1016/S0140-6736\(21\)02724-0](https://doi.org/10.1016/S0140-6736(21)02724-0).
3. Nakonieczna J; Wozniak A; Pieranski M; Rapacka-Zdonczyk A; Ogonowska P ; Grinholc M. Photoinactivation of ESKAPE pathogens: overview of novel therapeutic strategy. *Future Medicinal Chemistry.* **2019**, *11*(5), 443-461, <https://doi.org/10.4155/fmc-2018-0329>.
4. Kunhikannan S; Thomas CJ; Franks AE; Mahadevaiah S; Kumar S ; Petrovski S. Environmental hotspots for antibiotic resistance genes. *Microbiologyopen.* **2021**, *10*(3), e1197, <https://doi.org/10.1002/mbo3.1197>.
5. Monegro AF; Muppidi V ; Regunath H. Hospital acquired infections. StatPearls Publishing. <https://www.ncbi.nlm.nih.gov/books/NBK441857/>. Accessed on October 17, 2022.
6. Gernaey KV; van Loosdrecht MCM; Henze M; Lind M ; Jørgensen SB. Activated sludge wastewater treatment plant modelling and simulation: state of the art. *Environmental Modelling & Software.* **2004**, *19*(9), 763-783, <https://doi.org/10.1016/j.envsoft.2003.03.005>.

7. Leinonen R; Sugawara H; Shumway M ; on behalf of the International Nucleotide Sequence Database Collaboration. The Sequence Read Archive. *Nucleic Acids Research*. **2010**, *39*(suppl\_1), D19-D21, <https://doi.org/10.1093/nar/gkq1019>.
8. Chen S; Zhou Y; Chen Y ; Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. **2018**, *34*(17), i884-i890, <https://doi.org/10.1093/bioinformatics/bty560>.
9. Andrews S. FastQC: a quality control tool for high throughput sequence data. Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>. Accessed on October 26, 2021.
10. 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*. **2015**, *31*(10), 1674-1676, <https://doi.org/10.1093/bioinformatics/btv033>.
11. Wood DE ; Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biology*. **2014**, *15*(3), R46, <https://doi.org/10.1186/gb-2014-15-3-r46>.
12. Alcock BP; Raphenya AR; Lau TTY; Tsang KK; Bouchard M; Edalatmand A, et al. CARD 2020: antibiotic resistance surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res*. **2020**, *48*(D1), D517-d525, <https://doi.org/10.1093/nar/gkz935>.
13. Sridhar S; Worby C; Bronson R; Turbett S; Harris J; Ryan ET, et al. Alterations to the Gut Microbiomes and Acquisition of Bacteria Resistance Elements among US International Travelers. *Open Forum Infectious Diseases*. **2021**, *8*(Supplement\_1), S110-S110, <https://doi.org/10.1093/ofid/ofab466.180>.
14. Wan S; Xia M; Tao J; Pang Y; Yu F; Wu J, et al. Metagenomics Analysis Reveals the Microbial Communities, Antimicrobial Resistance Gene Diversity and Potential Pathogen Transmission Risk of Two Different Landfills in China. *Diversity*. **2021**, *13*(6), 230, <https://doi.org/10.3390/d13060230>.
15. Talat A; Blake KS; Dantas G ; Khan AU. Metagenomic Insight into Microbiome and Antibiotic Resistance Genes of High Clinical Concern in Urban and Rural Hospital Wastewater of Northern India Origin: a Major Reservoir of Antimicrobial Resistance. *Microbiol Spectr*. **2023**, e04102-04122, <https://doi.org/10.1128/spectrum.04102-22>.
16. Buchfink B; Reuter K ; Drost H-G. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nature Methods*. **2021**, *18*(4), 366-368, <https://doi.org/10.1038/s41592-021-01101-x>.
17. Li H. Seqtk. GitHub. <https://github.com/lh3/seqtk>. Accessed on October 30, 2021.
18. Bushnell B. BBtools. <https://sourceforge.net/projects/bbmap/>. Accessed on June 1, 2022.
19. Perry MR; Lepper HC; McNally L; Wee BA; Munk P; Warr A, et al. Secrets of the Hospital Underbelly: Patterns of Abundance of Antimicrobial Resistance Genes in Hospital Wastewater Vary by Specific Antimicrobial and Bacterial Family. *Front Microbiol*. **2021**, *12*, 703560, <https://doi.org/10.3389/fmicb.2021.703560>.
20. Abu-Rub LI; Johar A-RA; Al Mana H; Abdelrahman HA; Althani AA; Qotba H, et al. Bacterial indoor air contaminations in hospitals in MENA region: a systematic review. *International Journal of Environmental Health Research*. **2022**, <https://doi.org/10.1080/09603123.2022.2083087>.
21. Ali A; Sultan I; Mondal AH; Siddiqui MT; Gogry FA ; Haq QMR. Lentic and effluent water of Delhi-NCR: a reservoir of multidrug-resistant bacteria harbouring blaCTX-M, blaTEM and blaSHV type ESBL genes. *Journal of Water and Health*. **2021**, *19*(4), 592-603, <https://doi.org/10.2166/wh.2021.085>.
22. Sorgen A; Johnson J; Lambirth K; Clinton SM; Redmond M; Fodor A, et al. Characterization of Environmental and Cultivable Antibiotic-Resistant Microbial Communities Associated with Wastewater Treatment. *Antibiotics*. **2021**, *10*(4), 352, <https://doi.org/10.3390/antibiotics10040352>.
23. Fu P; Xu H; Jing C; Deng J; Wang H; Hua C, et al. Bacterial Epidemiology and Antimicrobial Resistance Profiles in Children Reported by the ISPED Program in China, 2016 to 2020. *Microbiol Spectr*. **2021**, *9*(3), e0028321, <https://doi.org/10.1128/Spectrum.00283-21>.
24. Chamoun K; Farah M; Araj G; Daoud Z; Moghnieh R; Salameh P, et al. Surveillance of antimicrobial resistance in Lebanese hospitals: retrospective nationwide compiled data. *International Journal of Infectious Diseases*. **2016**, *46*, 64-70, <https://doi.org/10.1016/j.ijid.2016.03.010>.
25. Hu F; Guo Y; Yang Y; Zheng Y; Wu S; Jiang X, et al. Resistance reported from China antimicrobial surveillance network (CHINET) in 2018. *European Journal of Clinical Microbiology & Infectious Diseases*. **2019**, *38*(12), 2275-2281, <https://doi.org/10.1007/s10096-019-03673-1>.
26. Azuma T; Usui M ; Hayashi T. Inactivation of Antibiotic-Resistant Bacteria in Wastewater by Ozone-Based Advanced Water Treatment

- Processes. *Antibiotics*. **2022**, *11*(2), 210, <https://doi.org/10.3390/antibiotics11020210>.
27. Ali OS; Hozayen WG; Almutairi AS; Edris SA; Abulfaraj AA; Ouf AA, et al. Metagenomic Analysis Reveals the Fate of Antibiotic Resistance Genes in a Full-Scale Wastewater Treatment Plant *Front Microbiol.* **2022**, *13*, <https://doi.org/10.3389/fmicb.2022.990736>.
29. Sękowska A. Raoultella spp.—clinical significance, infections and susceptibility to antibiotics. *Folia Microbiologica*. **2017**, *62*(3), 221-227, <https://doi.org/10.1007/s12223-016-0490-7>.
30. World Health Organization. WHO list of critically important antimicrobials for human medicine (WHO CIA list). World Health Organization. <https://apps.who.int/iris/bitstream/handle/10665/312266/9789241515528-eng.pdf>. Accessed on December 15, 2022.
31. Thaker M; Spanogiannopoulos P ; Wright GD. The tetracycline resistome. *Cellular and Molecular Life Sciences*. **2010**, *67*(3), 419-431, <https://doi.org/10.1007/s00018-009-0172-6>.
32. Maestre-Carballa L; Navarro-López V ; Martinez-Garcia M. A Resistome Roadmap: From the Human Body to Pristine Environments. *Front Microbiol.* **2022**, *13*, <https://doi.org/10.3389/fmicb.2022.858831>.
33. Yang Y; Li B; Ju F ; Zhang T. Exploring Variation of Antibiotic Resistance Genes in Activated Sludge over a Four-Year Period through a Metagenomic Approach. *Environmental Science & Technology*. **2013**, *47*(18), 10197-10205, <https://doi.org/10.1021/es4017365>.
- in Egypt. *Sustainability*. **2021**, *13*(20), 11131, <https://doi.org/10.3390/su132011131>.
28. Li A; Yan C; Zhang L; Liu S; Feng C; Zhang L, et al. Characterization and identification of a novel chromosomal class C  $\beta$ -lactamase, LAQ-1, and comparative genomic analysis of a multidrug resistance plasmid in *Lelliottia amnigena* P13.
34. Tao W; Zhang X-X; Zhao F; Huang K; Ma H; Wang Z, et al. High Levels of Antibiotic Resistance Genes and Their Correlations with Bacterial Community and Mobile Genetic Elements in Pharmaceutical Wastewater Treatment Bioreactors. *PLOS ONE*. **2016**, *11*(6), e0156854, <https://doi.org/10.1371/journal.pone.0156854>.
35. Nielsen TK; Browne PD ; Hansen LH. Antibiotic resistance genes are differentially mobilized according to resistance mechanism. *GigaScience*. **2022**, *11*, <https://doi.org/10.1093/gigascience/giac072>.
36. Piddock LJV. Multidrug-resistance efflux pumps ? not just for resistance. *Nature Reviews Microbiology*. **2006**, *4*(8), 629-636, <https://doi.org/10.1038/nrmicro1464>.
37. Uluseker C; Kaster KM; Thorsen K; Basiry D; Shobana S; Jain M, et al. A Review on Occurrence and Spread of Antibiotic Resistance in Wastewaters and in Wastewater Treatment Plants: Mechanisms and Perspectives. *Front Microbiol.* **2021**, *12*, <https://doi.org/10.3389/fmicb.2021.717809>.
38. Karkman A; Berglund F; Flach C-F; Kristiansson E ; Larsson DGJ. Predicting clinical resistance prevalence using sewage metagenomic data. *Communications Biology*. **2020**, *3*(1), 711, <https://doi.org/10.1038/s42003-020-01439-6>.