Meta-analysis reveals the global picture of antibiotic resistance gene prevalence across environments

Anna Abramova[1,2], Thomas U. Berendonk[3], Johan Bengtsson-Palme[1,2]*

[1] Department of Infectious Diseases, Institute of Biomedicine, The Sahlgrenska Academy, University of Gothenburg, Guldhedsgatan 10A, SE-413 46 Gothenburg, Sweden

[2] Centre for Antibiotic Resistance research (CARe) at University of Gothenburg, Gothenburg, Sweden

[3] Institute for Hydrobiology, Technische Universität Dresden, Dresden 01062, Germany

*Corresponding author: johan.bengtsson-palme@gu.se

Abstract

The environment is an important component in the emergence and transmission of antimicrobial resistance (AMR). Despite that, current AMR monitoring initiatives lack comprehensive reference data for the vast majority of environments. To enable monitoring to detect deviations from the normal background resistance levels in the environment, it is necessary to establish the typical baseline of AMR in a variety of settings. In an attempt to establish this baseline level, we here performed a comprehensive literature survey across 2516 scientific papers, 150 of which contained relevant qPCR data on antibiotic resistance genes (ARGs) in environments associated with potential routes of AMR dissemination. The collected data include 1487 samples distributed across 30 different countries and 17 environmental types, in a time span from 2005 to 2018. More than 330 different genes were identified from the collected studies. Most surveyed environments contained a diverse set of ARGs, but generally at low abundances. We used linear mixed models and

overrepresentation analysis to identify time trends and associations between ARGs and specific environments. Altogether these data represent a comprehensive overview of the occurrence and levels of ARGs in different environments, providing background data for risk assessment models within current and future AMR monitoring frameworks.

Keywords: antimicrobial resistance, AMR, qPCR, monitoring, surveillance

Introduction

According to the World Health Organization the increasing level of antibiotic resistance over the last decades is a considerable threat to global health[1]. Already, current estimates suggest that hundreds of thousands of deaths are caused by antibiotic resistant bacteria each year worldwide, and this number is expected to reach several millions annually within the next 30 years[2].

Antibiotic resistance has been mainly investigated in clinical and agricultural settings. However, the role of the environment as a source and dissemination route of resistance has been widely recognized in recent years[3-5]. Antibiotic resistance is widespread in nature and not limited to human-associated activities as evidenced by antibiotic resistance genes (ARGs) found in glaciers, permafrost and isolated caves which have virtually never been visited by humans[6-8]. It is even likely that resistance may be maintained independent of human activities in e.g. soil systems due to the presence of antibiotic-producing microorganisms, including certain fungal taxa[9]. Due to the extraordinary biological diversity in nature, the external environment can serve as a source of novel ARGs not yet present in clinically relevant pathogens[10]. Furthermore, the environment constitutes a dissemination route for resistant bacteria between humans and animals[11]. However, it is not yet clear to what extent the resistant environmental bacteria and the presence of ARGs in the environment contribute to acquisition and spread of resistance in clinically relevant bacteria. To address these questions, it is essential to identify and quantify resistance in the environment. This in turn would allow pinpointing which environments possess the highest risks to human and domestic animal health and would subsequently enable us to devise measures to prevent – or at least delay – recruitment and dissemination of resistance factors from the environment.

Monitoring schemes for antibiotic resistance have been developed and put to use in clinical and agricultural settings, but systematic monitoring for antimicrobial resistance (AMR) in the environment is still lacking[12]. Existing monitoring systems largely rely on culturing of pathogens and bacterial indicator species of interest[13], and these approaches may not be directly applicable to the monitoring of resistance in the environment. For example, the majority of environmental bacteria are not culturable using standard laboratory methods and the breakpoints for resistance are adapted for clinically relevant pathogens. Circumventing culturing drawbacks, quantitative real-time PCR (qPCR) allows detection and quantification of individual ARGs. However, a major hurdle for using qPCR for monitoring of AMR is the decision of which genes to select as targets in these qPCR assays. While large arrays that can measure the abundances of hundreds of ARGs in parallel exist[14], they are still fairly expensive and may not yet be ready for implementation in routine monitoring[15], and this would thus necessitate selection of a smaller number of highly relevant qPCR targets.

To address this problem, and to design efficient monitoring and surveillance strategies, there is a need for comprehensive background data on abundances and prevalence of ARGs occurring in both pristine and human-impacted environments. Establishing background levels of ARGs is a prerequisite for determining “safe levels” for ARGs naturally occurring in the environment and would also provide insights into to which extent an environment has been enriched with ARGs due to human activity. However, there is currently no, or very little, reference data available for the majority of environments and no agreed-upon set of genes to be used for environmental AMR monitoring.

The aim of this study was to perform a literature survey of studies containing qPCR data for ARGs in various environments and identify a set of genes and reference values for key environments for AMR transmission, which could be used for monitoring purposes. Through screening of more than 2500 scientific papers, we found that most surveyed environments contained a diverse set of ARGs, but generally at low abundances. However, it was clear

that the existing data on ARG abundances is unevenly distributed across the world and that there is a general lack of data for most environments and geographical areas. The study highlights the need for knowledge of environmental background levels of AMR to understand spatiotemporal changes in ARG abundances that can be associated with increased risks for adverse human health effects.

Materials and Methods

Data collection

Literature collection was performed by mining the PubMed database in April-May 2020, November 2020 and October 2021. The searches were performed using three strings: '"antibiotic resistance genes" AND (monitoring OR surveillance) AND qPCR AND', '"antibiotic resistance genes" AND "high throughput" AND' and '"antibiotic resistance" AND (qPCR OR "quantitative PCR") AND' complemented by environmental type: "sewage OR wastewater", "surface water", sediment, soil, food, "drinking water", "bathing water", "swimming pools", "bathing water", water, airport, ("public transport" OR bus OR train OR airplane OR boat OR tram OR underground OR metro), (industrial OR pollution), abandoned, cleanroom, animal and "one health". Environmental types were selected based on an assessment of potential transmission routes for antibiotic resistance[4,11]. Paper suitability was initially assessed by screening the title, abstract, and methods section. Only papers using qPCR as a quantification method and containing extractable data were selected. Review papers were not included in this study.

For each paper the following information was extracted: first author, DOI, year, country, sampling location, treatment/type (including distance from a contamination source, season,

type of fertilizer/antibiotic applied, stage of wastewater treatment process, source; e.g. for "water" it would be "potable", "reclaimed", "river" etc.), sample type, and abundance of antibiotic resistance genes (ARGs). For several studies missing information about the exact year of data collection, we extrapolated the sampling year based on the average time required from the data collection to the publication inferred from the other studies, corresponding to an estimate of two years. The ARG abundances were collected from text, tables and figures. Information from figures was extracted using WebPlotDigitizer v4.3 (https://automeris.io/WebPlotDigitizer). Abundances were reported as log10 ARG abundance per 16S rRNA. ARGs with abundance below detection limit were not included in the analysis.

Data analysis

Data were analyzed in RStudio v1.3.1073[16] with the R package ggplot2 v3.3.5[17] used to plot the graphs. To determine the abundance distributions for ARGs across environments, density plots of the ARG abundances were used, and the R summary function was used to find the quartiles for each ARG. To investigate if there were any changes in ARGs levels over time, we built a linear mixed model (LMM). LMM allows incorporation of not only fixed effects but also random effects to account for cluster-correlated data from distinct sources of variability[18]. The model was designed with ARG *Abundance* as a response variable and sampling *Year* as explanatory variable and included sampling *Country* as a covariate:

$$Abundance = Year + (1|Country) + \varepsilon$$

Analyses were performed using lme4 v1.1-27.1[19] and lmerTest v3.1-3[20] packages.

To identify genes associated with particular environments, we used the R package indicspecies v1.7.9. This analysis uses relative frequency and abundance of species (in our

152 case genes were treated as “species”) in defined groups, to identify those that are the most
153 characteristic of each group[21]. Only genes found in more than 20 samples were included in
154 this analysis. The function *multipatt* was run with 9999 permutations. The results were
155 visualized as a network using Cytoscape (cytoscape.org).
156
157 To identify a set of suitable qPCR targets we investigated how well the data from collected
158 genes could predict the abundance of the other genes in the data matrix by performing a
159 correlation analysis. Only genes with more than 10 entries were included in the analysis. R
160 package stats (version 3.6.2) *cor* function with “pairwise.complete.obs” was used to calculate
161 Spearman correlations.

162

163 Results and Discussion

164

165 **Current data on AMR in the environment is geographically biased**

166 Mining of the PubMed database resulted in 2516 papers and revealed a bias in the existing
167 literature towards already fairly well studied environments (Figure S1). Unsurprisingly, the
168 majority of the identified papers matched environmental categories that have been already
169 extensively studied as primary sources of antibiotics and/or antibiotic resistance to the
170 environment. These include environments linked to wastewater treatment plants, hospitals
171 and industrial facilities, as well as agriculture and livestock production. This also highlights
172 the undersampled nature of the other environments. To bring an example, it is suggested
173 that human mobility, in particular international travel, could contribute to dissemination of
174 ARGs across the globe[10,22-24]. Despite that, categories associated with travel activity, such as
175 “public transport” and “airports”, are represented by just a handful of papers. Similarly, water
176 associated with recreational activities is another poorly investigated environmental category,

177 despite that accidental ingestion of water during, for example, swimming is considered to be
178 a potential route for resistant environmental bacteria into the human gut[25].
179
180 In total, we identified 150 studies (1487 samples) containing relevant information on levels of
181 ARGs measured by qPCR. These studies covered 17 different types of environments with
182 “water” and “effluent” being the most commonly investigated (covered by 46 and 33 studies,
183 respectively). In contrast, “food” and “air” were mentioned in just five studies or less. The
184 samples collected in these studies originated from 30 different countries (Figure 1A), with
185 China being the most studied among them, both in terms of number of studies and samples.
186 The distribution of studies highlights white spots on the map; in particular, there were no
187 studies originating from South America and just a handful of studies from Africa. This is
188 worrisome, as South America and Africa are among the regions with the highest rates of
189 self-medication and supply of non-prescribed antibiotics in the world[26,27]. In 2014, a WHO
190 report outlined Africa’s lack of established AMR surveillance systems[1] and a review from
191 2017 reported that more than a third of the countries on the continent did not have recent
192 AMR data published in the public domain, and that only little of the reported data were from
193 surveillance[28]. This scarcity of publicly available data on AMR is corroborated by the current
194 literature survey. The majority of samples in the selected studies were collected between
195 2015-2018 with a gradual increase in the number of publications over the last years (Figure
196 1C).
197

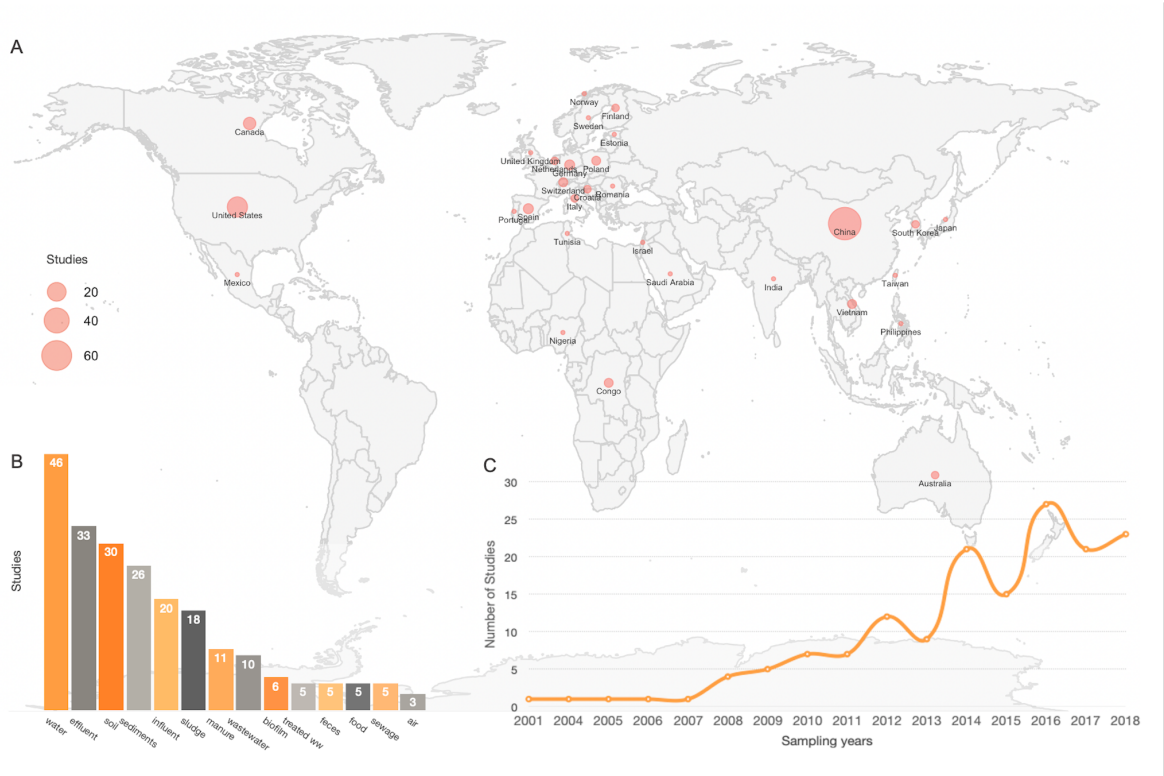


Figure 1. Overview of the collected data. Studies distributed by A) country, B) sampled environment, and C) year.

Environmental ARGs are diverse in almost all environments

We identified 331 different genes that were quantified in at least one of the collected studies. They belonged to several different ARG classes with β-lactamase and multi-drug resistance genes being the most diverse classes (54 and 46 unique ARGs, respectively) (Table S1). When ARG data were stratified by country (Figure 2), it was evident that overall the diversity of targeted and detected ARGs was fairly low (86% of the studies targeted less than 10 ARGs on average). Most of the studies used singleplex qPCR (119) and 29 used high throughput qPCR (HT-qPCR).

Most of the studied environments were characterized by a diverse set of investigated ARG classes (Figure S2), but despite that a limited but obvious bias was observed towards certain classes of ARGs. In particular, tetracycline and β-lactam resistance genes were among the

most commonly used targets for qPCR studies. Interestingly, studies investigating sewage targeted mostly β-lactamases in comparison to the other classes. This may be due to these genes being perceived to be associated with a relatively higher risk to human health, which would justify intensive research on their removal in sewage treatment plants. A range of macrolide-lincosamide-streptogramin B (MLSB) resistance genes were looked for and reported from all the investigated environments. Multidrug resistance genes showed a similar pattern with the exception of being absent from “feces” and “sewage” environments, which is most likely due to a limited number of studies targeting these particular ARG classes rather than this being their actual distribution. Colistin resistance genes were reported only from five environments in studies published starting from 2018. This class is an interesting case, since colistin belongs to the family of polymyxins which used to be one of the last classes of antibiotics for which resistance was not plasmid-mediated. In 2015, the first plasmid-borne colistin resistance gene was described[29], and since then it has been reported in many bacterial species and samples across the world, some predating the initial discovery[30,31], although it was not often targeted in the studies we analyzed. There were also several classes which were profiled consistently across all the environments, including sulfonamides (with the *sul1* and *sul2* genes reported from all the environments), vancomycin and aminoglycoside resistance genes.

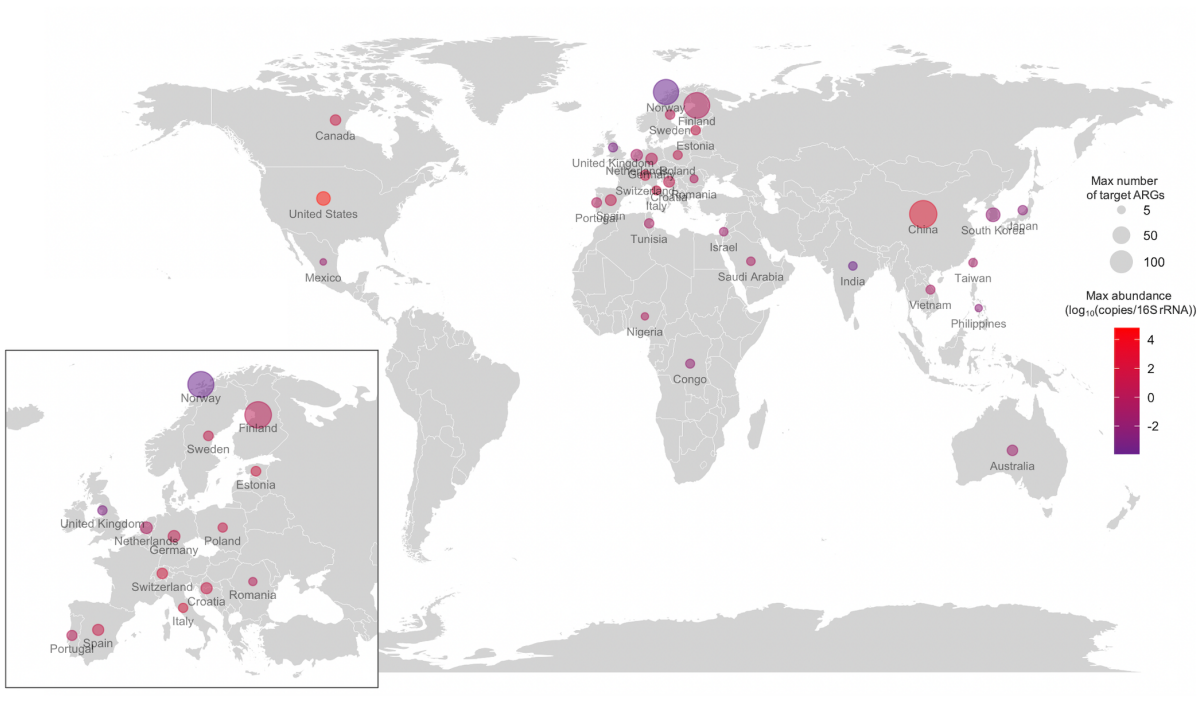


Figure 2. Maximum diversity and maximum abundance of ARGs in sampled countries.

Around one in a thousand environmental bacteria carry clinical ARGs

Most of the ARGs present in more than 20 studies were characterized by higher than average abundance (Figure 3) and were, unsurprisingly, widespread across the studied environments (Figure 4). Among the 20 most reported genes, tetracycline resistance genes were the predominant class of ARGs. Sulfonamide resistance genes *sul1* and *sul2* were the two most reported genes, with *sul1* reported in 68% of studies and from all the environments. The *intI1* gene was also reported from all the investigated environments and was characterized by the highest average abundance among all the genes in our data matrix. These genes are among the ones suggested as good monitoring targets by Berendonk et al.[4]. Genes encoding extended-spectrum β-lactamases, including $bla_{CTX\text{-}M}$, bla_{TEM} and bla_{SHV}, were among the common targets in qPCR studies, which is not surprising since they are ARGs commonly encountered in clinically relevant pathogens. The plasmid-mediated fluoroquinolone resistance gene *qnrS* has been reported from aquatic environments and wildlife[4,32,33]. A study of the distribution of this gene in wildlife and water

252 reservoirs suggested that it can be horizontally transferred even in the absence of a
253 selective pressure from antibiotics[34]. The most commonly reported genes also included the
254 erythromycin resistance genes *ermB* and *ermF,* and the aminoglycoside resistance genes
255 *strB* and *aadA1*. Interestingly, there were several genes that were often highly abundant
256 when detected, but which were not often reported, making them potentially suitable
257 additional targets for AMR monitoring (*vanA, tnpA4, tnpA5, sul3, qacEdelta, mexF, IS613,*
258 *floR, ereA* and *aadA2*).
259
260 For each ARG with more than a single observation, we estimated its typical environmental
261 abundance levels (Figure S3). We found that for most ARGs, the typically reported
262 abundances fell in an interval from 10^{-5} to 10^{-3} copies per 16S rRNA, roughly corresponding
263 to one ARG copy in a thousand bacteria (Table S2). That said, some ARGs, such as the
264 *mcr1* colistin resistance gene (mostly below 10^{-5} ARGs/16S), had lower typical abundances,
265 and tetracycline resistance genes *tetO* and *tetQ* often occurred in abundances above 1 copy
266 per 16S rRNA in fecal samples (Figure S3). Interestingly, overall *environmental* abundances
267 of these two *tet* genes were in the typical 10^{-5} to 10^{-3} range, indicating that these genes may
268 be good indicators of pollution with human fecal matter. They may even be more suitable
269 than *intI1*, which has been proposed as a marker of human anthropogenic impact in the
270 past[35] but consistently appeared in 10^{-5} to 10^{-3} copies per 16S rRNA in most environments in
271 our data (Figure S3).
272
273 Taken together, it seems normal to encounter most ARGs in abundances between 10^{-5} to
274 10^{-3} copies per 16S rRNA in a given environment, although we refer the reader to Figure S3
275 and Table S2 for specific values for individual ARGs. In general, we advise the use of 10
276 times the third quartile (3Q) values for a given ARG (Table S2) to determine the upper limit
277 for what should be viewed as a deviation from normal background levels in the environment.
278 This should be combined with studying the overall and environment-specific abundance
279 distributions for the same gene (Figure S3). It is worth noting that this analysis does not

consider whether the detected ARG abundances are the result of anthropogenic activities[36,37] or natural selection for ARGs in these environments[9,38].

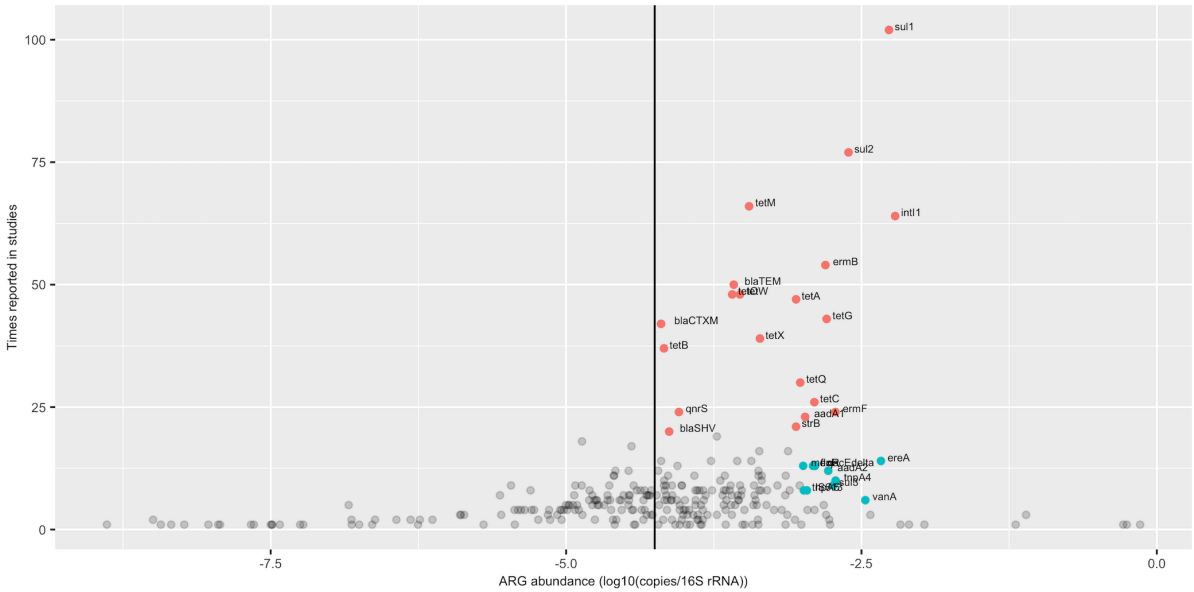


Figure 3. Abundance of most reported ARGs. ARGs reported in more than 20 studies are represented in orange, and ARGs reported less but which were generally highly abundant are shown in blue. The line represents an average abundance across all data points.

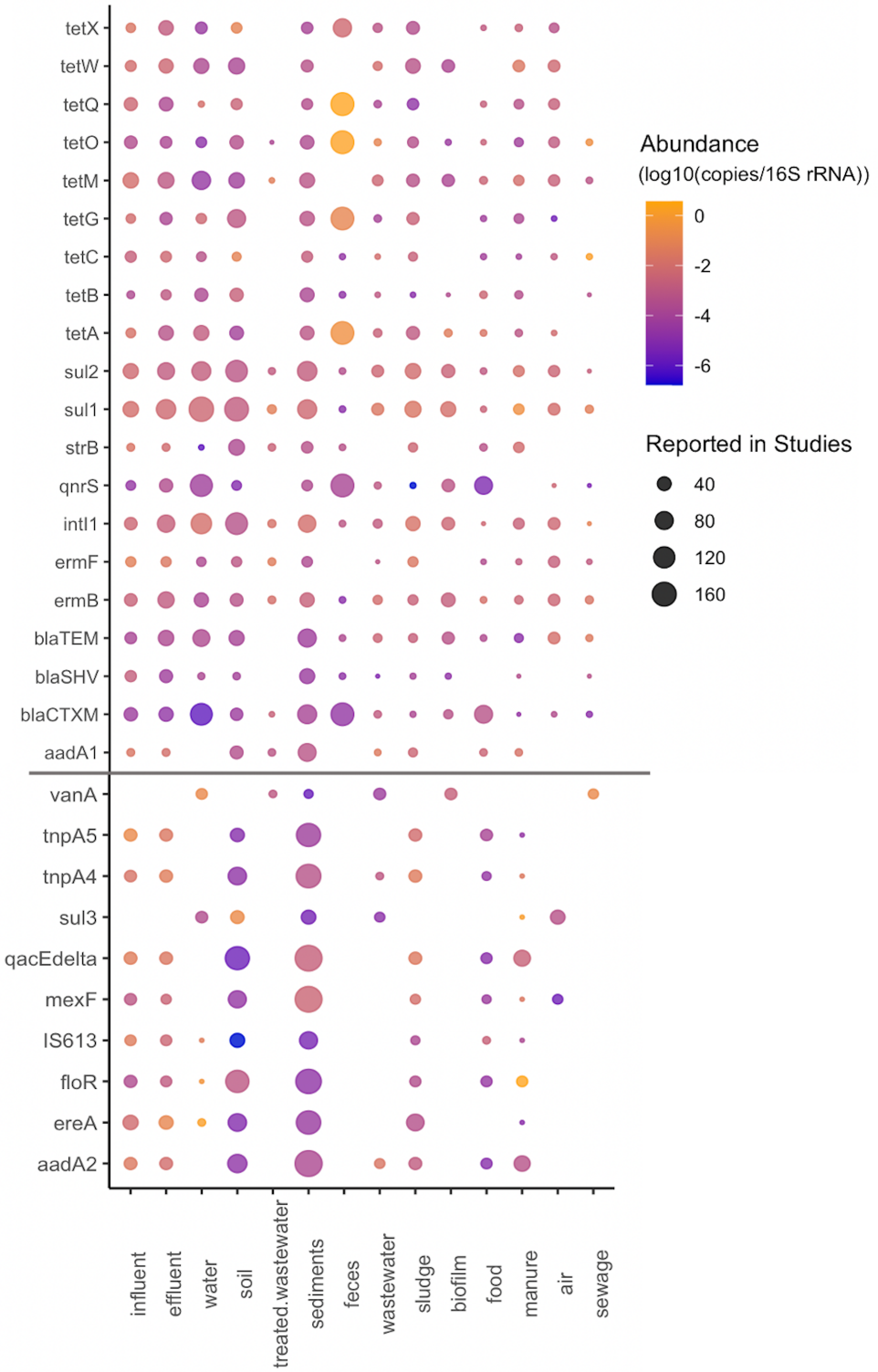


Figure 4. Distribution of the most reported and abundant ARGs across environments. Above the line are ARGs reported in more than 20 studies and below are the ARGs reported less but highly abundant.

It is unclear if ARG abundances have increased over time

To investigate if there were changes in ARG abundances over time, we used linear mixed models on the data for all of the ARGs in each of the environments, accounting for variability between countries (Figure S5). Due to the increased sensitivity of the qPCR methods and inclusion of a larger number of ARGs profiled in each study over time, we chose to use linear mixed models for the maximum values rather than all ARG values reported. Using this approach, we found positive trends for most of the environments (for cases where there was enough data to perform the estimates), except for in sediments, sludge, wastewater and treated wastewater where the trends were negative (Table S3). That said, the trends were significant only for soil (increase) and treated wastewater (decrease). Thus, the most prominent finding of this analysis might be that in most cases there was too little data to draw any conclusions on ARG abundance changes over time, highlighting the need for time series data to understand the long-term development of antibiotic resistance in the environment.

Most ARGs are not associated with specific environments

We performed an overrepresentation analysis showing which ARGs were associated with particular environments (Figure S6, Table S4). We found that a substantial number of ARGs were specifically associated with sludge, sewage, manure and soil, but not the other environments. However, comparably few ARGs were specifically associated with just a single environment. Some notable examples of such single-environment associations include *mcr1* in feces, *qnrD* in wastewater, *strA*, *strB*, *tnpA3* and *aphA3* in manure, and *aadA3* in sediments.

Abundance data for some ARGs provide redundant information

The set of most reported genes consists of potential candidates to be used as qPCR targets, since they were on average more abundant and were also found in most environments (Figures 3 and 4). To explore if they would be good predictors of overall ARG abundances,

320 we performed a correlation analysis of the most reported genes and the rest of the genes in
321 the data matrix (Figure S4). Interestingly, this analysis revealed a certain degree of
322 redundancy in terms of what information can be gained from the proposed monitoring
323 targets. ARGs proposed as monitoring targets that strongly correlate with each other convey
324 similar information on ARG abundance and diversity and may therefore not be very useful to
325 use in the same panel for ARG monitoring in the environment. Genes that are overall
326 redundant are the *tetO*, *tetX* and *tetA* genes, *strB* and *ermF*, $bla_{CTX\text{-}M}$ and bla_{SHV}, *tetG* and
327 *tetM*, as well as *intI1* and *sul2* (Figure S4). If resources are constrained, it would seem wise
328 not to use several genes in these smaller groups together in the same ARG panel for
329 monitoring.
330

331 **Priorities for environmental monitoring of AMR**

332 It is well established that antibiotic resistance is an ancient phenomenon, and that the
333 environment represents a rich reservoir of resistance factors which could be transferred to
334 human pathogens under the right conditions[9,38-40]. Therefore, environmental monitoring of
335 antibiotic resistance is important in order to protect human and animal health. However, a
336 systematic framework for monitoring in the environment is still lacking[12], and defining *what,*
337 *where* and *how* to monitor is crucial for such efforts to contribute relevant data[4]. There are
338 several main objectives for monitoring of AMR in the environment, including *i)* assessment of
339 risks for transmission of already resistant pathogens to humans via environmental routes, *ii)*
340 analysis of general levels of resistance in the human population, and *iii)* monitoring the
341 evolution of resistance in the environment to identify emerging resistance threats[41].
342

343 **Risk ranking of ARGs**

344 There have been several attempts to create a system for ranking ARGs by risk to human
345 health. Two common criteria used to estimate risks include known presence of an ARG in
346 human pathogens and whether it is located on a mobile genetic element, such as a plasmid.
347 For instance, Martinez et al.[42] proposed a ranking system which assigns the highest rank to

genes that have already been shown to be a clinical problem and the lowest rank to genes whose resistance function was inferred solely by homology. However, ARGs that are well-known and have been around for decades in clinical pathogens may be less of a concern than those that have recently emerged and that could potentially pose the next big threat, but are not yet widespread among pathogens[10]. Thus, ARGs which are not ubiquitous and not yet present in human pathogens must be carefully considered when assessing risks associated with environmental resistomes. A recent paper by Zhang et al.[43] presented another approach based on microbial ecology, with the decisive ranking criteria being enrichment of the ARG by anthropogenic activities. ARGs enriched by anthropogenic activities and correlated with antibiotics contamination should, by this line of thinking, pose a higher risk than those that are not enriched[43]. This is, in theory, a sound approach; however, a major issue would be determining which ARGs that would be enriched by antibiotic exposure, as this is in most cases not known outside of ARGs commonly occurring in human pathogens. Another factor to consider in risk ranking could be the presence of an ARG in a variety of different environments[44] or in different bacterial species, including potential human pathogens, which would suggest it has a high capacity for horizontal gene transfer and transmission[45]. From a human health risk perspective, it may be particularly relevant to target ARGs that confer resistance to last resort antibiotics, such as carbapenems, vancomycin or colistin[4]. Many of these genes are currently not abundant in environmental samples, which may present a challenge for environmental monitoring[13]. It has also been proposed to monitor the diversity of ARGs, as a high ARG diversity would be associated with a larger set of genes that could potentially be transferred to pathogens[13,46]. Along the same lines, it is worth considering how well a certain ARG would predict the abundances of other ARGs in an environment, as it has been shown that some ARGs have better predictive power than others[46].

374 **Monitoring targets identified in the current study**

375 We compared the list of genes identified in the present literature survey with proposed
376 candidates for environmental monitoring from earlier studies (Table S5). Some of the
377 identified ARGs are already common in routine clinical monitoring (e.g.
378 https://www.oxfordbiosystems.com/AMR, https://www.opgen.com/acuitas/amr-gene-panel/),
379 including clinically relevant β-lactamases such as $bla_{CTX\text{-}M}$, bla_{TEM} and bla_{SHV}, the
380 sulfonamide resistance gene *sul1,* the vancomycin resistance gene *vanA,* as well as *ermB*
381 and *tetM* conferring resistance to macrolides and tetracyclines, respectively. These genes
382 are already common both in the environment and among human pathogens, and therefore
383 their usefulness in environmental monitoring is somewhat questionable. At the same time,
384 they may function as a gauge of the total antibiotic resistance content in an environment, as
385 some of them have been shown to be good predictors of the total resistome, particularly *tetM*
386 and bla_{TEM}[46].
387
388 Considering the most common ARGs in our survey that are not yet included in routine
389 clinical monitoring (Table 1) highlighted additional future monitoring targets, including the
390 class 1 integrase *intI1*, which is already a widely used marker for anthropogenic AMR
391 dissemination[35]. Despite that this gene is not a resistance gene itself, is often linked to ARGs
392 and according to the existing literature it is widespread in the environment, found in both
393 pathogenic and commensal bacteria, enriched by anthropogenic activities and correlates
394 well with the abundances of other commonly occurring ARGs. This gene fulfills most of the
395 criteria for being a suitable monitoring target, except perhaps that it is already so widespread
396 (Figure 3; Figure S3) that it may not be able to sufficiently capture minor or long-term
397 changes in the environment over time and space.
398
399 Similarly, the sulfonamide resistance gene *sul2* has also been suggested as a monitoring
400 target by Berendonk et al. (2015)[4] and was among the most reported in our data and is
401 widespread across all the environments. However, *sul2* added little information compared to

402 *intI1* (Figure S4), so these two genes would be mostly redundant in AMR monitoring. In
403 comparison, *sul3* was not so often reported despite being rather abundant in the cases when
404 it was looked for. According to previous studies, *sul3* is typically less frequent and abundant
405 than *sul1* and *sul2*[47,48]. However, it has been shown to be associated with class 1 integrons
406 and has been detected on a conjugative plasmid, suggesting its potential to be horizontally
407 transferred and disseminated in the environment[49]. Interestingly, *sul3* was first detected in an
408 *E. coli* isolate from a pig, and it was later found in both healthy and diseased humans.
409 Furthermore, it is enriched in polluted environments and is present in both commensal and
410 pathogenic bacteria[50]. Overall, these three *sul* genes are good indicators of mobile antibiotic
411 resistance, but may suffer from not being that predictive for total ARG abundance and
412 diversity[46], and their value for monitoring in addition to *intI1* is somewhat doubtful.
413
414 The plasmid-mediated fluoroquinolone resistance gene *qnrS* was among the most reported
415 genes in our study. It was found in all environments except manure and treated wastewater.
416 As it is the only commonly reported fluoroquinolone ARG, it is an interesting target for
417 monitoring. However, it has the drawback of generally not being able to induce clinically
418 relevant levels of resistance without additional resistance mechanisms, which makes its
419 clinical impact limited.
420
421 The list of most reported genes was dominated by tetracycline resistance genes. Among
422 them, *tetM* was the third most often reported gene in the investigated studies, and has been
423 previously proposed as a monitoring target and included in clinical screening[4,46]. Tetracycline
424 resistance genes are among the most powerful ARGs for predicting the diversity and
425 abundance of other ARGs[46] and therefore, despite their widespread distribution, these genes
426 can still be useful in monitoring to extrapolate other parts of the resistome.
427
428

429 **Table 1. Characterization of genes identified in the current survey as potential targets for**

430 **environmental AMR monitoring**

431

ARG	Ecology	Association with MGEs	Presence in ESKAPE pathogens	Wide niche adaptation	Distribution across environments	Abundance variability (0,01-1,95)	Correlation coefficient and number of significant correlations
intI1	Enriched in polluted environments[35]	class 1 integron-integrase gene	yes[51]	yes[51]	All environments	1,61	0,64 (97)
sul2	Correlates with concentrations of antibiotics [52,53]	yes[54]	yes[54]	yes[55]	All environments	1,56	0,80 (150)
sul3	Correlates with heavy metals pollution[56]	yes[57]	?	yes[58]	6	1,54	0,71 (8)
qnrS	Enriched in fluoroquinolone contaminated environment [59,60]	yes[61]	yes[62]	yes[63]	12	1,56	0,68 (23)
strB	Correlates with heavy metals pollution[64]	yes[65]	yes[66]	yes[67]	10	1,47	0,64 (83)
ermF	Correlates with antibiotics level[68]	yes[69]	?	yes [70, 71]	12	1,44	0,66 (68)
tetB	Correlates with antibiotics level[52]	yes[72]	yes[73]	yes[74]	12	0,29	0,73 (98)
tetC	Correlates with antibiotics levels[75,76]	yes[77]	yes[78]	yes[79]	12	1,49	0,70 (23)
tetX	Correlates with antibiotics[80]	yes[81]	yes[82]	yes[83]	11	0,76	0,67 (83)
tetQ	Correlates with heavy metals[84]	yes[85]	?	yes[86]	11	1,35	0,72 (76)
tetW	Correlated with the concentrations of tetracycline residues[87]	yes[88]	?	yes	10	0,63	0,69 (60)
aadA1	Correlates with pollution levels[89]	yes	yes	yes	9	1,22	0,63 (97)
aadA2	Correlates with pollution levels[89]	yes	yes	yes	8	0,90	0,61 (88)
vanA	Correlates with antibiotic residues[90]	yes[91]	yes[92]	yes[91]	6	1,34	0,72 (3)
floR	Correlates with heavy metals[89]	yes	yes[93]	yes	8	1,62	0,70 (80)
ereA	Correlates with pollution[94]	yes[95]	yes	yes	7	1,36	0,63 (85)
mexF	Correlates with AB use[96]	yes[97]	yes[98]	yes[99]	8	1,10	0,77 (100)
qacEdelta	Correlates with heavy metals[100]	yes[101]	yes[102]	?	7	1,46	0,80 (134)

432

433

434 There were also several genes that were not often included in qPCR studies, but when they
435 were, they often appeared in relatively high abundances. Such genes could be interesting
436 additional monitoring targets in future AMR monitoring schemes. Among these genes was
437 the vancomycin resistance gene *vanA,* a previously suggested indicator of antibiotic
438 resistance contamination of clinical origin which is thought to be uncommon in the
439 environment[90]. In our data, however, *vanA* has been about as abundant as other ARGs in
440 the environment when it has been looked for. Another gene that seems to correlate with
441 anthropogenic activities is *ereA*, a macrolide resistance gene which has previously been
442 reported to be the most abundant in metal polluted soil[103] and is enriched by long-term
443 application of manure[94]. In contrast, the *mexF* gene has been reported from many different
444 environments, including soil, sediments and water[22] and is suggested to naturally occur in
445 unaffected environments such as pristine Antarctic soils[104, 105], indicating that it might be a
446 useful target for identifying enrichment of environmental ARGs.
447
448 **Specific uses of qPCR as a monitoring tool**
449 In this study, we have exclusively focused on the use of qPCR as a tool for monitoring the
450 abundance of ARGs in the environment. There are, however, other ways of measuring the
451 levels of AMR, including selective culturing and metagenomic sequencing approaches. In
452 this context, it is worth considering not only the appropriate targets for monitoring, but also
453 which methods that give the most informative output in a cost-efficient manner. In general,
454 the qPCR technique has several benefits, but also some drawbacks as a surveillance tool.
455 An important advantage is that it is a highly sensitive method that can detect much lower
456 levels of ARGs than is currently possible using shotgun metagenomic sequencing[106].
457 Furthermore, it can operate on very minute quantities of DNA, which makes it suitable also
458 for low-biomass samples. In addition, as it operates on DNA, it will be able to detect ARGs in
459 non-culturable bacteria and can function also on complex samples with many different
460 species. On the other hand, qPCR does not provide any information about the genetic
461 context of ARGs, such as whether they are chromosomal or mobile or if they are co-located

with other ARGs. Similarly, it does not indicate which hosts carry a given ARG. Crucially, it is also restricted to a predetermined set of target ARGs (or other types of genes) which limits the reusability of the data. For example, when the *mcr1* colistin resistance gene was discovered in 2015[107], it very quickly was detected in archived metagenomic sequence data[24], but such 'retrospective monitoring' would not be possible to perform using qPCR datasets. The limiting factor of predefined targets can be partially overcome by using qPCR arrays with hundreds of genes[15], but this also comes with increased costs and may not be feasible for large-scale routine monitoring of environmental AMR. In addition, the technique would say close to nothing about emerging ARGs that may become the next generation resistance threats in human pathogens[10], which makes qPCR-based monitoring practically useless as an early-warning system for emerging resistance, unless a panel of potential future ARGs is screened for. That said, qPCR clearly has its uses in monitoring for transmission of known high-risk ARGs via the environment, where it can provide relatively quick data on specific deviations from the normal background levels.

A more fundamental aspect of such deviations is whether an increase in the abundance of ARGs or mobile genetic elements in a particular environment or at a specific time point is a relevant indicator of a selective pressure for resistance. Importantly, an increase of ARG abundances without any context is more likely to be an indicator of human pollution[36], but it cannot be ruled out that such a change could be due to a specific selection pressure from antibiotics[108], by co-selection from other antibacterial compounds[109], or simply from taxonomic shifts that are unrelated to antibiotic resistance. In the end, the suitability of qPCR for environmental AMR surveillance comes down to the purpose of monitoring and what type of actions one might want to take based on the monitoring results. If transmission of known high-risk ARGs is the sole purpose of monitoring, qPCR is an excellent method thanks to its sensitivity. However, if identification of emergent resistance threats is the goal, qPCR is unlikely to give useful guidance; instead, shotgun metagenomics or selective culturing followed with genetic profiling would provide more useful information.

490 **Outcomes and recommendations**

491 In this study, we performed a literature survey to explore the abundance and prevalence of
492 ARGs in various environments as quantified by qPCR. We found that, overall, previous
493 suggestions for ARGs to be included in environmental AMR monitoring[4,46] seem relevant.
494 Particularly, inclusion of the *intI1*, *sul1*, bla_{TEM}, $bla_{CTX\text{-}M}$ and *qnrS* genes in environmental
495 monitoring seems essential, along with a selection of the tetracycline genes *tetM* (but not
496 *tetG*), *tetQ*, *tetW*, *tetC*, and either of *tetA* or *tetO*. However, there are also genes that are not
497 often looked for in qPCR surveys that perhaps should be, including *sul3*, *vanA*, *tetH*, *aadA2*,
498 *floR*, *ereA* and *mexF*. These genes are abundant in some environments, but were not often
499 included in qPCR studies of environmental AMR. We also provide environmental baseline
500 levels for the ARGs studied through qPCR (Figure 5; Figure S3); for most ARGs the typical
501 abundance falls in an interval from 10^{-5} to 10^{-3}. It should be noted that this is the range of
502 normal abundances of ARGs, and should not be considered a maximum acceptable limit of
503 these ARGs in any given environment. Such maximum acceptable levels need to be
504 determined taking risks to human health (and potentially the environment) into account[5,110],
505 and would also need to consider transmission routes to humans[11]. The absence of clear
506 trends of increases or decreases in ARG abundances over time indicates a need for more
507 systematic time series data on ARG abundances in a variety of environments. Our results
508 also highlight the scarcity of AMR data from parts of the world, particularly from Africa and
509 South America, and underscores the need for a concerted effort to quantify typical
510 background levels of AMR in the environment more broadly to enable efficient environmental
511 AMR surveillance schemes akin to those that exist in clinical settings.

512

513

514

515

Acknowledgments

JBP acknowledges funding from the Swedish Research Council (VR; grant 2019-00299) under the frame of JPI AMR (EMBARK; JPIAMR2019-109), the Centre for Antibiotic Resistance Research at the University of Gothenburg, the Sahlgrenska Academy at the University of Gothenburg, and the Swedish Cancer and Allergy fund (Cancer- och Allergifonden). TUB acknowledges funding of the JPI AMR – EMBARK project funded by the Bundesministerium für Bildung, und Forschung (BMBF) under grant number F01KI1909A. We would like to thank Emil Burman, Sebastian Wettersten and Tora Hulterström for constructive feedback on the manuscript.

References

1. WHO, *Antimicrobial resistance: global report on surveillance*. 2014: World Health Organization.
2. O'Neill, J., *Antimicrobial resistance.* Tackling a crisis for the health and wealth of nations, 2014.
3. Bengtsson-Palme, J., *Assessment and management of risks associated with antibiotic resistance in the environment*, in *Management of Emerging Public Health Issues and Risks*. 2019, Elsevier. p. 243-263.
4. Berendonk, T.U., et al., *Tackling antibiotic resistance: the environmental framework.* Nature Reviews Microbiology, 2015. **13**(5): p. 310-317.
5. Larsson, D.J., et al., *Critical knowledge gaps and research needs related to the environmental dimensions of antibiotic resistance.* Environment international, 2018. **117**: p. 132-138.
6. Mindlin, S. and M. Petrova, *On the origin and distribution of antibiotic resistance: permafrost bacteria studies.* Molecular Genetics, Microbiology and Virology, 2017. **32**(4): p. 169-179.
7. Pawlowski, A.C., et al., *A diverse intrinsic antibiotic resistome from a cave bacterium.* Nature communications, 2016. **7**(1): p. 1-10.
8. Van Goethem, M.W., et al., *A reservoir of 'historical'antibiotic resistance genes in remote pristine Antarctic soils.* Microbiome, 2018. **6**(1): p. 1-12.
9. Bahram, M., et al., *Structure and function of the global topsoil microbiome.* Nature, 2018. **560**(7717): p. 233-237.
10. Bengtsson-Palme, J. and D.G.J. Larsson, *Antibiotic resistance genes in the environment: prioritizing risks.* Nature Reviews Microbiology, 2015. **13**(6): p. 396-396.
11. Bengtsson-Palme, J., E. Kristiansson, and D.G.J. Larsson, *Environmental factors influencing the development and spread of antibiotic resistance.* FEMS microbiology reviews, 2018. **42**(1): p. fux053.
12. Pruden, A., et al., *Seizing the moment: now is the time for integrated global surveillance of antimicrobial resistance in wastewater environments.* Curr Opin Microbiol, 2021. **64**: p. 91-99.
13. Vikesland, P.J., et al., *Toward a comprehensive strategy to mitigate dissemination of environmental sources of antibiotic resistance*. 2017, ACS Publications.

14. Stedtfeld, R.D., et al., *Primer set 2.0 for highly parallel qPCR array targeting antibiotic resistance genes and mobile genetic elements.* FEMS Microbiology Ecology, 2018. **94**(9).
15. Waseem, H., et al., *Contributions and Challenges of High Throughput qPCR for Determining Antimicrobial Resistance in the Environment: A Critical Review.* Molecules, 2019. **24**(1): p. 163.
16. R Team, *RStudio: integrated development environment for R. RStudio, PBC, Boston*. 2020.
17. Wickham, H., *ggplot2-Elegant Graphics for Data Analysis. Springer International Publishing.* Cham, Switzerland, 2016.
18. Harrison, X.A., et al., *A brief introduction to mixed effects modelling and multi-model inference in ecology.* PeerJ, 2018. **6**: p. e4794.
19. Bates, D., et al., *Fitting linear mixed-effects models using lme4.* arXiv preprint arXiv:1406.5823, 2014.
20. Kuznetsova, A., P.B. Brockhoff, and R.H. Christensen, *lmerTest package: tests in linear mixed effects models.* Journal of statistical software, 2017. **82**(1): p. 1-26.
21. Cáceres, M.D. and P. Legendre, *Associations between species and groups of sites: indices and statistical inference.* Ecology, 2009. **90**(12): p. 3566-3574.
22. D'Souza, A.W., et al., *Destination shapes antibiotic resistance gene acquisitions, abundance increases, and diversity changes in Dutch travelers.* Genome medicine, 2021. **13**(1): p. 1-21.
23. Frost, I., et al., *Global geographic trends in antimicrobial resistance: the role of international travel.* Journal of travel medicine, 2019. **26**(8): p. taz036.
24. Von Wintersdorff, C.J., et al., *Dissemination of antimicrobial resistance in microbial ecosystems through horizontal gene transfer.* Frontiers in microbiology, 2016. **7**: p. 173.
25. Leonard, A.F., et al., *Exposure to and colonisation by antibiotic-resistant E. coli in UK coastal water users: Environmental surveillance, exposure assessment, and epidemiological study (Beach Bum Survey).* Environment international, 2018. **114**: p. 326-333.
26. Auta, A., et al., *Global access to antibiotics without prescription in community pharmacies: a systematic review and meta-analysis.* Journal of Infection, 2019. **78**(1): p. 8-18.
27. Yeika, E.V., et al., *Comparative assessment of the prevalence, practices and factors associated with self-medication with antibiotics in Africa.* Tropical Medicine & International Health, 2021. **26**(8): p. 862-881.
28. Tadesse, B.T., et al., *Antimicrobial resistance in Africa: a systematic review.* BMC infectious diseases, 2017. **17**(1): p. 1-17.
29. Liu, Y.Y., et al., *Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study.* Lancet Infect Dis, 2016. **16**(2): p. 161-8.
30. Hu, Y., et al., *Dissemination of the mcr-1 colistin resistance gene.* The Lancet Infectious Diseases, 2016. **16**(2): p. 146-147.
31. Ruppé, E., et al., *Dissemination of the mcr-1 colistin resistance gene.* The Lancet Infectious Diseases, 2016. **16**(3): p. 290-291.
32. Literak, I., et al., *Antibiotic-resistant Escherichia coli bacteria, including strains with genes encoding the extended-spectrum beta-lactamase and QnrS, in waterbirds on the Baltic Sea Coast of Poland.* Applied and environmental microbiology, 2010. **76**(24): p. 8126-8134.
33. Xiong, W., et al., *Antibiotics, antibiotic resistance genes, and bacterial community composition in fresh water aquaculture environment in China.* Microbial ecology, 2015. **70**(2): p. 425-432.
34. Vredenburg, J., et al., *Quinolone-resistant E scherichia coli isolated from birds of prey in P ortugal are genetically distinct from those isolated from water environments*

612 *and gulls in P ortugal, S pain and S weden.* Environmental microbiology, 2014. **16**(4):
613 p. 995-1004.
614 35. Gillings, M.R., et al., *Using the class 1 integron-integrase gene as a proxy for*
615 *anthropogenic pollution.* The ISME journal, 2015. **9**(6): p. 1269-1279.
616 36. Karkman, A., K. Pärnänen, and D.G.J. Larsson, *Fecal pollution can explain antibiotic*
617 *resistance gene abundances in anthropogenically impacted environments.* Nature
618 communications, 2019. **10**(1): p. 1-8.
619 37. Pruden, A., et al., *Antibiotic resistance genes as emerging contaminants: studies in*
620 *northern Colorado.* Environmental science & technology, 2006. **40**(23): p. 7445-7450.
621 38. D'Costa, V.M., et al., *Antibiotic resistance is ancient.* Nature, 2011. **477**(7365): p.
622 457-461.
623 39. Bengtsson-Palme, J., V. Jonsson, and S. Hess, *What is the role of the environment*
624 *in the emergence of novel antibiotic resistance genes?-A modelling approach.*
625 Environmental Science & Technology, 2021. **55**(23): p. 15734–15743.
626 40. Forsberg, K.J., et al., *The shared antibiotic resistome of soil bacteria and human*
627 *pathogens.* science, 2012. **337**(6098): p. 1107-1111.
628 41. Huijbers, P.M.C., C.F. Flach, and D.G.J. Larsson, *A conceptual framework for the*
629 *environmental surveillance of antibiotics and antibiotic resistance.* Environ Int, 2019.
630 **130**: p. 104880.
631 42. Martinez, J.L., T.M. Coque, and F. Baquero, *Prioritizing risks of antibiotic resistance*
632 *genes in all metagenomes.* Nat Rev Microbiol, 2015. **13**(6): p. 396.
633 43. Zhang, A.-N., et al., *An omics-based framework for assessing the health risk of*
634 *antimicrobial resistance genes.* Nature communications, 2021. **12**(1): p. 1-11.
635 44. Manaia, C.M., et al., *Antibiotic resistance in urban aquatic environments: can it be*
636 *controlled?* Applied microbiology and biotechnology, 2016. **100**(4): p. 1543-1557.
637 45. Ishii, S., *Quantification of antibiotic resistance genes for environmental monitoring:*
638 *current methods and future directions.* Current Opinion in Environmental Science &
639 Health, 2020. **16**: p. 47-53.
640 46. Bengtsson-Palme, J., *The diversity of uncharacterized antibiotic resistance genes*
641 *can be predicted from known gene variants—but not always.* Microbiome, 2018. **6**(1):
642 p. 125.
643 47. Hoa, P.T.P., et al., *Detection of the sul1, sul2, and sul3 genes in sulfonamide-*
644 *resistant bacteria from wastewater and shrimp ponds of north Vietnam.* Science of
645 the Total Environment, 2008. **405**(1-3): p. 377-384.
646 48. Suzuki, S., et al., *Contamination of antibiotics and sul and tet (M) genes in veterinary*
647 *wastewater, river, and coastal sea in Thailand.* Science of The Total Environment,
648 2021. **791**: p. 148423.
649 49. Antunes, P., J. Machado, and L. Peixe, *Dissemination of sul3-containing elements*
650 *linked to class 1 integrons with an unusual 3' conserved sequence region among*
651 *Salmonella isolates.* Antimicrobial agents and chemotherapy, 2007. **51**(4): p. 1545-
652 1548.
653 50. Hammerum, A.M., et al., *Detection of sul1, sul2 and sul3 in sulphonamide resistant*
654 *Escherichia coli isolates obtained from healthy humans, pork and pigs in Denmark.*
655 International journal of food microbiology, 2006. **106**(2): p. 235-237.
656 51. Zhang, A.N., et al., *Conserved phylogenetic distribution and limited antibiotic*
657 *resistance of class 1 integrons revealed by assessing the bacterial genome and*
658 *plasmid collection.* Microbiome, 2018. **6**(1): p. 130.
659 52. Leng, Y., et al., *Tetracyclines, sulfonamides and quinolones and their corresponding*
660 *resistance genes in coastal areas of Beibu Gulf, China.* Science of The Total
661 Environment, 2020. **714**: p. 136899.
662 53. Xu, Y., et al., *Occurrence and distribution of antibiotics, antibiotic resistance genes in*
663 *the urban rivers in Beijing, China.* Environmental pollution, 2016. **213**: p. 833-840.
664 54. Hamidian, M., S.J. Ambrose, and R.M. Hall, *A large conjugative Acinetobacter*
665 *baumannii plasmid carrying the sul2 sulphonamide and strAB streptomycin*
666 *resistance genes.* Plasmid, 2016. **87**: p. 43-50.

55. Matter, D., et al., *Small multidrug resistance plasmids in Actinobacillus porcitonsillarum.* Plasmid, 2008. **59**(2): p. 144-152.
56. Xu, Y., et al., *Effect of the selective pressure of sub-lethal level of heavy metals on the fate and distribution of ARGs in the catchment scale.* Environmental Pollution, 2017. **220**: p. 900-908.
57. Wu, S., et al., *Prevalence and characterization of plasmids carrying sulfonamide resistance genes among Escherichia coli from pigs, pig carcasses and human.* Acta Veterinaria Scandinavica, 2010. **52**(1): p. 1-7.
58. Byrne-Bailey, K., et al., *Prevalence of sulfonamide resistance genes in bacterial isolates from manured agricultural soils and pig slurry in the United Kingdom.* Antimicrobial Agents and Chemotherapy, 2009. **53**(2): p. 696-702.
59. Castrignanò, E., et al., *Enantiomeric profiling of quinolones and quinolones resistance gene qnrS in European wastewaters.* Water research, 2020. **175**: p. 115653.
60. Wang, L., et al., *Macrolide-and quinolone-resistant bacteria and resistance genes as indicators of antibiotic resistance gene contamination in farmland soil with manure application.* Ecological Indicators, 2019. **106**: p. 105456.
61. Poirel, L., et al., *Plasmid-mediated quinolone resistance determinant qnrS in Enterobacter cloacae.* Clinical microbiology and infection, 2006. **12**(10): p. 1021-1023.
62. Al Shaikhli, N.H., et al., *Antibiotic resistance determinants of Acinetobacter baumannii strains isolated from nosocomial infections.* 2020.
63. Pallecchi, L., et al., *High prevalence of qnr genes in commensal enterobacteria from healthy children in Peru and Bolivia.* Antimicrobial agents and chemotherapy, 2009. **53**(6): p. 2632-2635.
64. Xue, J., et al., *Occurrence of heavy metals, antibiotics, and antibiotic resistance genes in different kinds of land-applied manure in China.* Environmental Science and Pollution Research, 2021. **28**(29): p. 40011-40021.
65. Tauch, A., et al., *The 79,370-bp conjugative plasmid pB4 consists of an IncP-1β backbone loaded with a chromate resistance transposon, the strA-strB streptomycin resistance gene pair, the oxacillinase gene bla NPS-1, and a tripartite antibiotic efflux system of the resistance-nodulation-division family.* Molecular Genetics and Genomics, 2003. **268**(5): p. 570-584.
66. Samuelsen, Ø., et al., *Molecular epidemiology of metallo-β-lactamase-producing Pseudomonas aeruginosa isolates from Norway and Sweden shows import of international clones and local clonal expansion.* Antimicrobial agents and chemotherapy, 2010. **54**(1): p. 346-352.
67. Sundin, G. and C. Bender, *Dissemination of the strA-strB streptomycin-resistance genes among commensal and pathogenic bacteria from humans, animals, and plants.* Molecular ecology, 1996. **5**(1): p. 133-143.
68. Yi, X., M. Wang, and Z. Zhou, *The potential impact of naturally produced antibiotics, environmental factors, and anthropogenic pressure on the occurrence of erm genes in urban soils.* Environmental Pollution, 2019. **245**: p. 282-289.
69. Chung, W.O., et al., *Mobile elements carrying ermF and tetQ genes in Gram-positive and Gram-negative bacteria.* Journal of Antimicrobial Chemotherapy, 1999. **44**(3): p. 329-335.
70. Luo, H., et al., *Identification of ribosomal RNA methyltransferase gene erm F in Riemerella anatipestifer.* Avian Pathology, 2015. **44**(3): p. 162-168.
71. Milaković, M., et al., *Characterization of macrolide resistance in bacteria isolated from macrolide-polluted and unpolluted river sediments and clinical sources in Croatia.* Science of The Total Environment, 2020. **749**: p. 142357.
72. Han, J.E., et al., *Plasmid mediated tetracycline resistance of Vibrio parahaemolyticus associated with acute hepatopancreatic necrosis disease (AHPND) in shrimps.* Aquaculture Reports, 2015. **2**: p. 17-21.

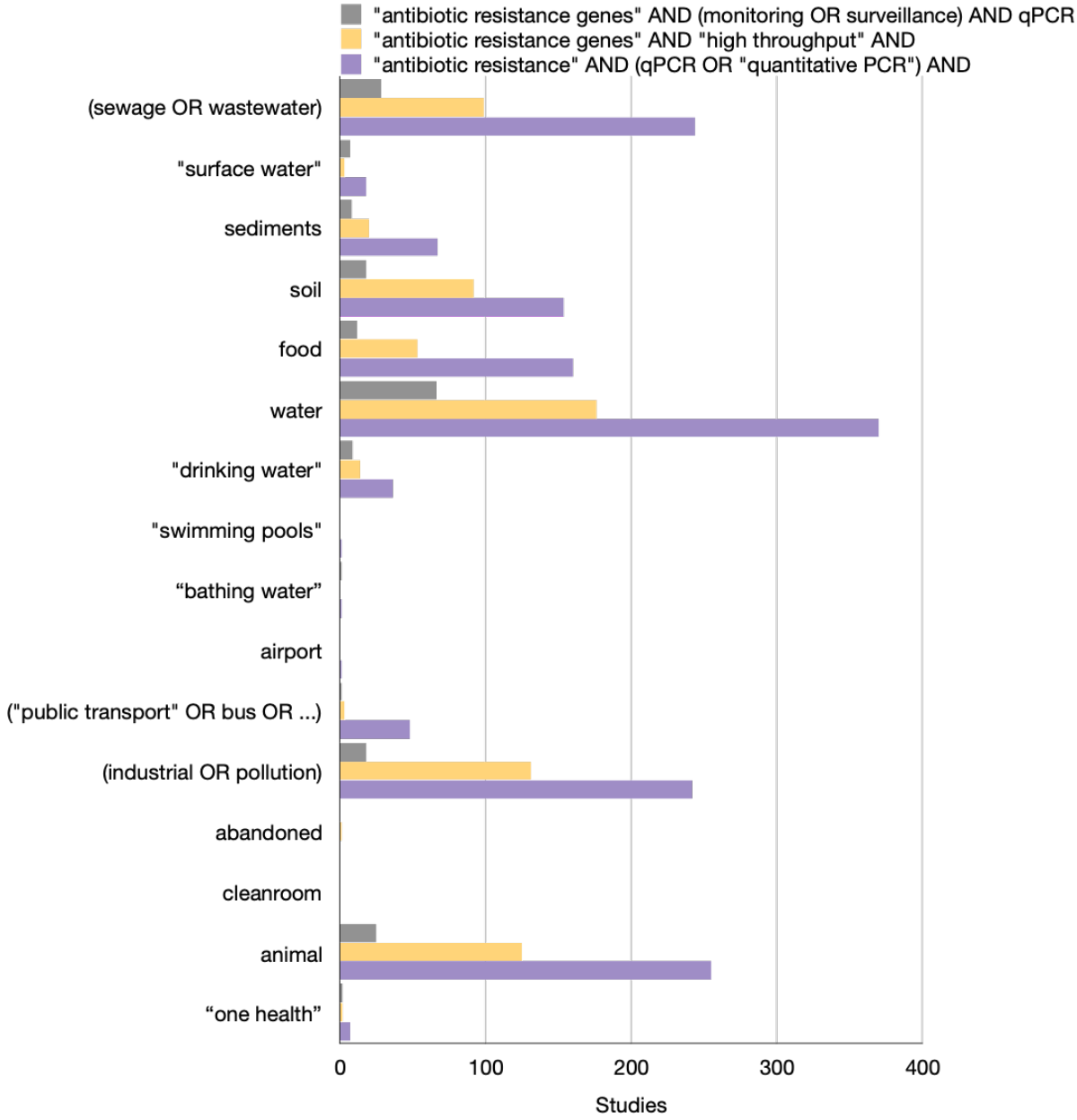
721 73. Ragupathi, N.K.D., et al., *Plasmid profiles among some ESKAPE pathogens in a*
722 *tertiary care centre in south India.* The Indian journal of medical research, 2019.
723 **149**(2): p. 222.
724 74. Fairchild, A., et al., *Effects of orally administered tetracycline on the intestinal*
725 *community structure of chickens and on tet determinant carriage by commensal*
726 *bacteria and Campylobacter jejuni.* Applied and environmental microbiology, 2005.
727 **71**(10): p. 5865-5872.
728 75. Chen, B., et al., *Differentiating anthropogenic impacts on ARGs in the Pearl River*
729 *Estuary by using suitable gene indicators.* Water research, 2013. **47**(8): p. 2811-
730 2820.
731 76. Huang, Y.-H., et al., *Occurrence and distribution of antibiotics and antibiotic resistant*
732 *genes in water and sediments of urban rivers with black-odor water in Guangzhou,*
733 *South China.* Science of the total environment, 2019. **670**: p. 170-180.
734 77. Shi, Y., et al., *Potential dissemination mechanism of the tetC gene in Aeromonas*
735 *media from the aerobic biofilm reactor under oxytetracycline stresses.* Journal of
736 Environmental Sciences, 2021. **105**: p. 90-99.
737 78. Isichei-Ukah, B. and O. Enabulele, *Analysis of Class 1 Integrons and Antibiotic*
738 *Resistance Genes in Pseudomonas aeruginosa Strains from Benin City, Nigeria.*
739 Journal of Applied Sciences and Environmental Management, 2020. **24**(4): p. 633-
740 637.
741 79. Das, A., et al., *Circulation of oxytetracycline-and ciprofloxacin-resistant commensal*
742 *Escherichia coli strains in broiler chickens and farm environments, Bangladesh.*
743 Veterinary World, 2020. **13**(11): p. 2395.
744 80. Duan, M., et al., *Factors that affect the occurrence and distribution of antibiotic*
745 *resistance genes in soils from livestock and poultry farms.* Ecotoxicology and
746 environmental safety, 2019. **180**: p. 114-122.
747 81. Li, R., et al., *Emergence of Plasmid-Mediated Resistance Genes tet (X) and mcr-1 in*
748 *Escherichia coli Clinical Isolates from Pakistan.* Msphere, 2021. **6**(4): p. e00695-21.
749 82. Markley, J.L. and T.A. Wencewicz, *Tetracycline-inactivating enzymes.* Frontiers in
750 microbiology, 2018. **9**: p. 1058.
751 83. Bryan, A., N. Shapir, and M.J. Sadowsky, *Frequency and distribution of tetracycline*
752 *resistance genes in genetically diverse, nonselected, and nonclinical Escherichia coli*
753 *strains isolated from diverse human and animal sources.* Applied and environmental
754 microbiology, 2004. **70**(4): p. 2503-2507.
755 84. Lu, L., et al., *Antibiotic resistance gene abundances associated with heavy metals*
756 *and antibiotics in the sediments of Changshou Lake in the three Gorges Reservoir*
757 *area, China.* Ecological Indicators, 2020. **113**: p. 106275.
758 85. Tribble, G.D., et al., *Genetic analysis of mobile tetQ elements in oral Prevotella*
759 *species.* Anaerobe, 2010. **16**(6): p. 604-609.
760 86. Veloo, A., et al., *Prevalence of antimicrobial resistance genes in Bacteroides spp.*
761 *and Prevotella spp. Dutch clinical isolates.* Clinical Microbiology and Infection, 2019.
762 **25**(9): p. 1156. e9-1156. e13.
763 87. Wu, N., et al., *Abundance and diversity of tetracycline resistance genes in soils*
764 *adjacent to representative swine feedlots in China.* Environmental Science &
765 Technology, 2010. **44**(18): p. 6933-6939.
766 88. Rozman, V., et al., *Characterization of antimicrobial resistance in lactobacilli and*
767 *bifidobacteria used as probiotics or starter cultures based on integration of*
768 *phenotypic and in silico data.* International journal of food microbiology, 2020. **314**: p.
769 108388.
770 89. Zhou, Q., et al., *Dissemination of resistance genes in duck/fish polyculture ponds in*
771 *Guangdong Province: correlations between Cu and Zn and antibiotic resistance*
772 *genes.* Environmental Science and Pollution Research, 2019. **26**(8): p. 8182-8193.
773 90. Narciso-da-Rocha, C., et al., *blaTEM and vanA as indicator genes of antibiotic*
774 *resistance contamination in a hospital–urban wastewater treatment plant system.*
775 Journal of global antimicrobial resistance, 2014. **2**(4): p. 309-315.

91. Werner, G., et al., *Host range of enterococcal vanA plasmids among Gram-positive intestinal bacteria.* Journal of Antimicrobial Chemotherapy, 2011. **66**(2): p. 273-282.
92. Navidinia, M., et al., *Molecular characterization of resistance genes in MDR-ESKAPE pathogens.* J Pure Appl Microbiol, 2017. **11**(2): p. 779-92.
93. Cloeckaert, A., S. Baucheron, and E. Chaslus-Dancla, *Nonenzymatic chloramphenicol resistance mediated by IncC plasmid R55 is encoded by a floR gene variant.* Antimicrobial Agents and Chemotherapy, 2001. **45**(8): p. 2381-2382.
94. Li, J., et al., *Long-term manure application increased the levels of antibiotics and antibiotic resistance genes in a greenhouse soil.* Applied soil ecology, 2017. **121**: p. 193-200.
95. Matsuoka, M., et al., *Characteristic expression of three genes, msr (A), mph (C) and erm (Y), that confer resistance to macrolide antibiotics on Staphylococcus aureus.* FEMS Microbiology Letters, 2003. **220**(2): p. 287-293.
96. Han, L., et al., *Development of antibiotic resistance genes in soils with ten successive treatments of chlortetracycline and ciprofloxacin.* Environmental Pollution, 2019. **253**: p. 152-160.
97. Norberg, P., et al., *The IncP-1 plasmid backbone adapts to different host bacterial species and evolves through homologous recombination.* Nature Communications, 2011. **2**(1): p. 268.
98. Li, X.-Z., N. Barré, and K. Poole, *Influence of the MexA-MexB-OprM multidrug efflux system on expression of the MexC-MexD-OprJ and MexE-MexF-OprN multidrug efflux systems in Pseudomonas aeruginosa.* Journal of Antimicrobial Chemotherapy, 2000. **46**(6): p. 885-893.
99. Binns, A.N. and J. Zhao, *The MexE/MexF/AmeC efflux pump of Agrobacterium tumefaciens and its role in Ti plasmid virulence gene expression.* Journal of bacteriology, 2020. **202**(8): p. e00609-19.
100. Buta, M., et al., *Sewage sludge in agriculture–the effects of selected chemical pollutants and emerging genetic resistance determinants on the quality of soil and crops–a review.* Ecotoxicology and Environmental Safety, 2021. **214**: p. 112070.
101. Gillings, M., et al., *The evolution of class 1 integrons and the rise of antibiotic resistance.* Journal of bacteriology, 2008. **190**(14): p. 5095-5100.
102. Surleac, M., et al., *Whole genome sequencing snapshot of multi-drug resistant Klebsiella pneumoniae strains from hospitals and receiving wastewater treatment plants in Southern Romania.* PloS one, 2020. **15**(1): p. e0228079.
103. Chen, J., et al., *Bacterial heavy-metal and antibiotic resistance genes in a copper tailing dam area in northern China.* Frontiers in microbiology, 2019. **10**: p. 1916.
104. Stedtfeld, R.D., et al., *Antimicrobial resistance dashboard application for mapping environmental occurrence and resistant pathogens.* FEMS microbiology ecology, 2016. **92**(3).
105. Wang, F., et al., *Influence of soil characteristics and proximity to Antarctic research stations on abundance of antibiotic resistance genes in soils.* Environmental science & technology, 2016. **50**(23): p. 12621-12629.
106. Lundström, S.V., et al., *Minimal selective concentrations of tetracycline in complex aquatic bacterial biofilms.* Science of the Total Environment, 2016. **553**: p. 587-595.
107. Liu, Y.-Y., et al., *Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study.* The Lancet infectious diseases, 2016. **16**(2): p. 161-168.
108. Bengtsson-Palme, J., et al., *Shotgun metagenomics reveals a wide array of antibiotic resistance genes and mobile elements in a polluted lake in India.* Frontiers in Microbiology, 2014. **5**.
109. Pal, C., et al., *Co-occurrence of resistance genes to antibiotics, biocides and metals reveals novel insights into their co-selection potential.* BMC Genomics, 2015. **16**(1): p. 964.

110. Pruden, A., et al., *Management Options for Reducing the Release of Antibiotics and Antibiotic Resistance Genes to the Environment.* Environmental Health Perspectives, 2013. **121**(8): p. 878-885.

Supplementary Figures and Tables

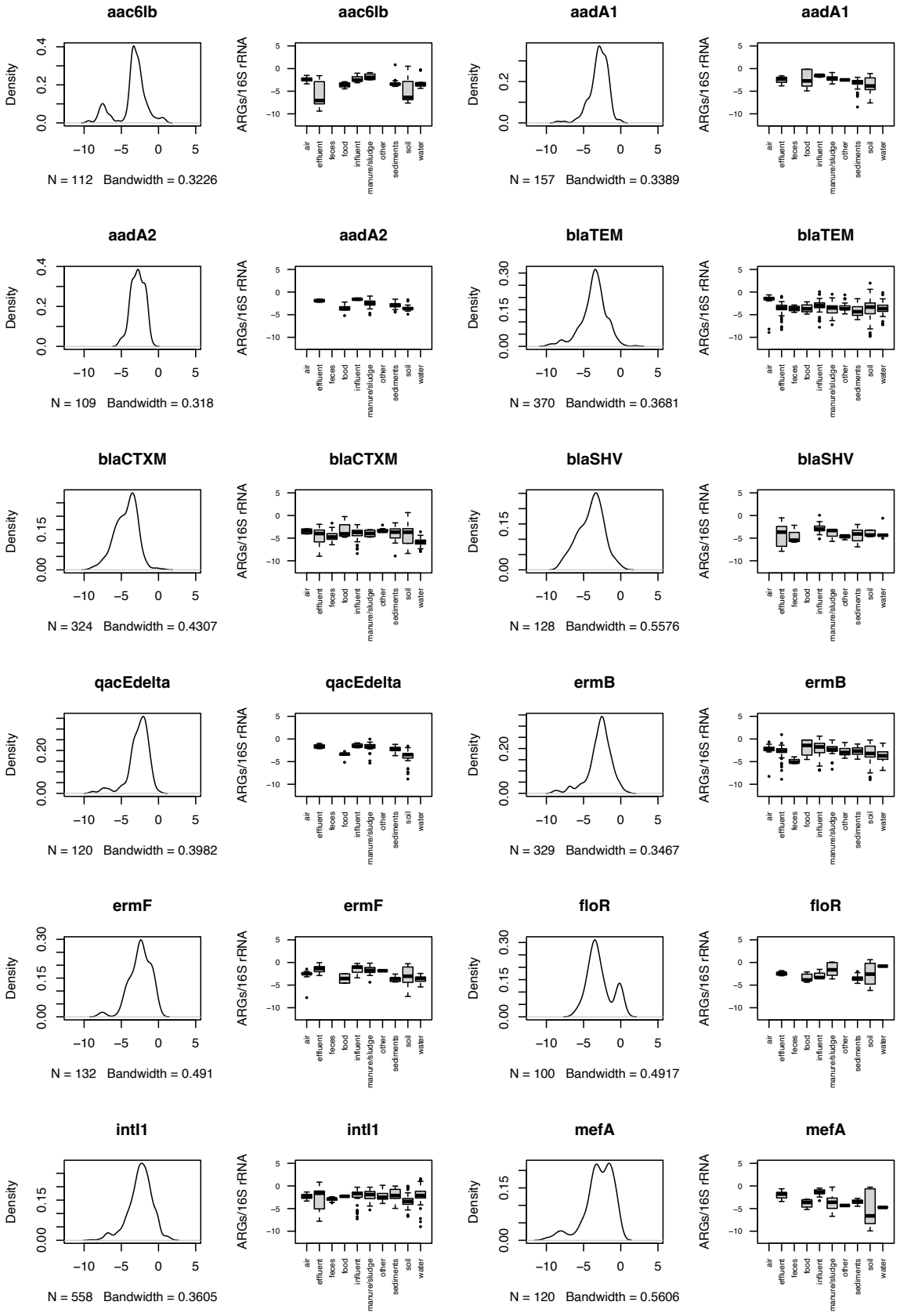
Figure S1. Total number of collected studies per each environmental category.

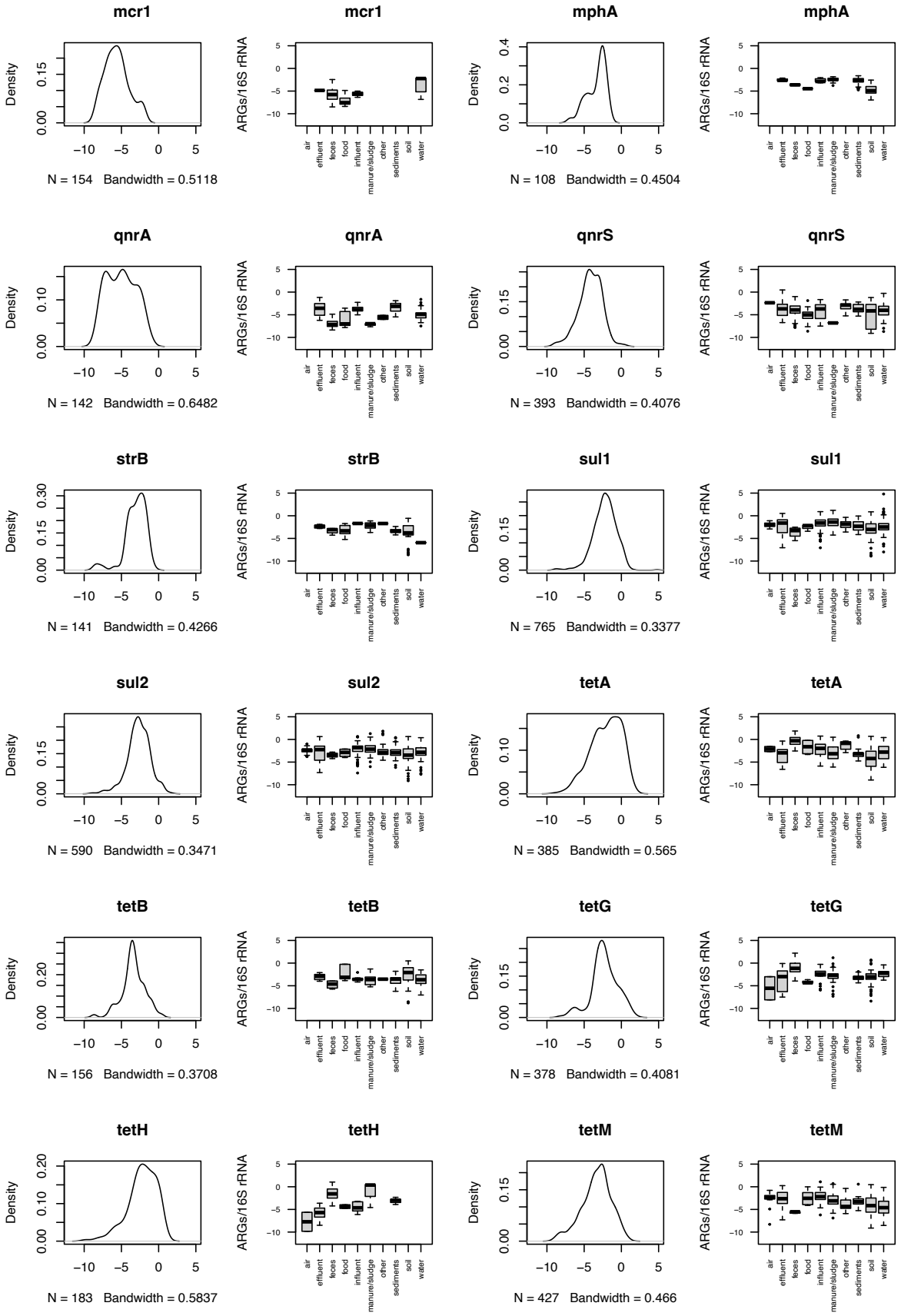


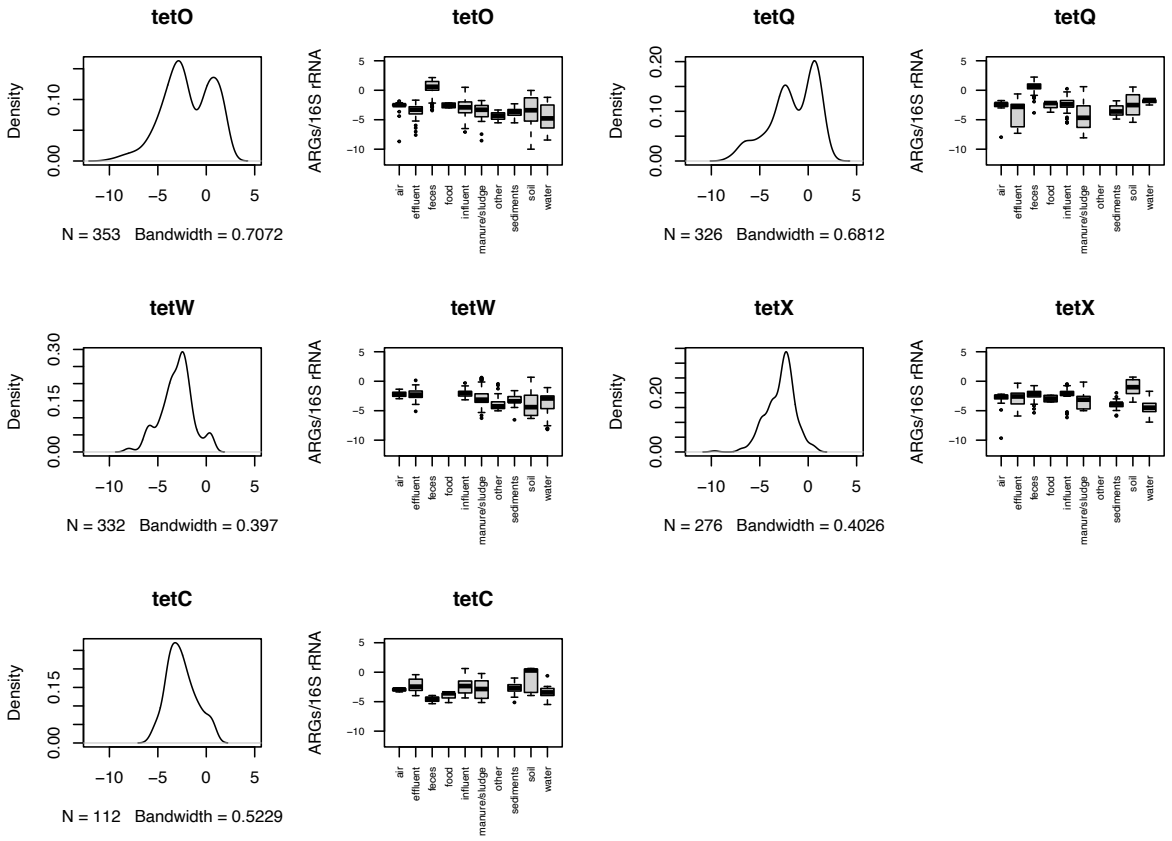
842 **Figure S2. Proportion of ARGs in each class reported in different environments.**

843
844
845
846

Figure S3. The distribution of abundances for the ARGs with more than 100 observations

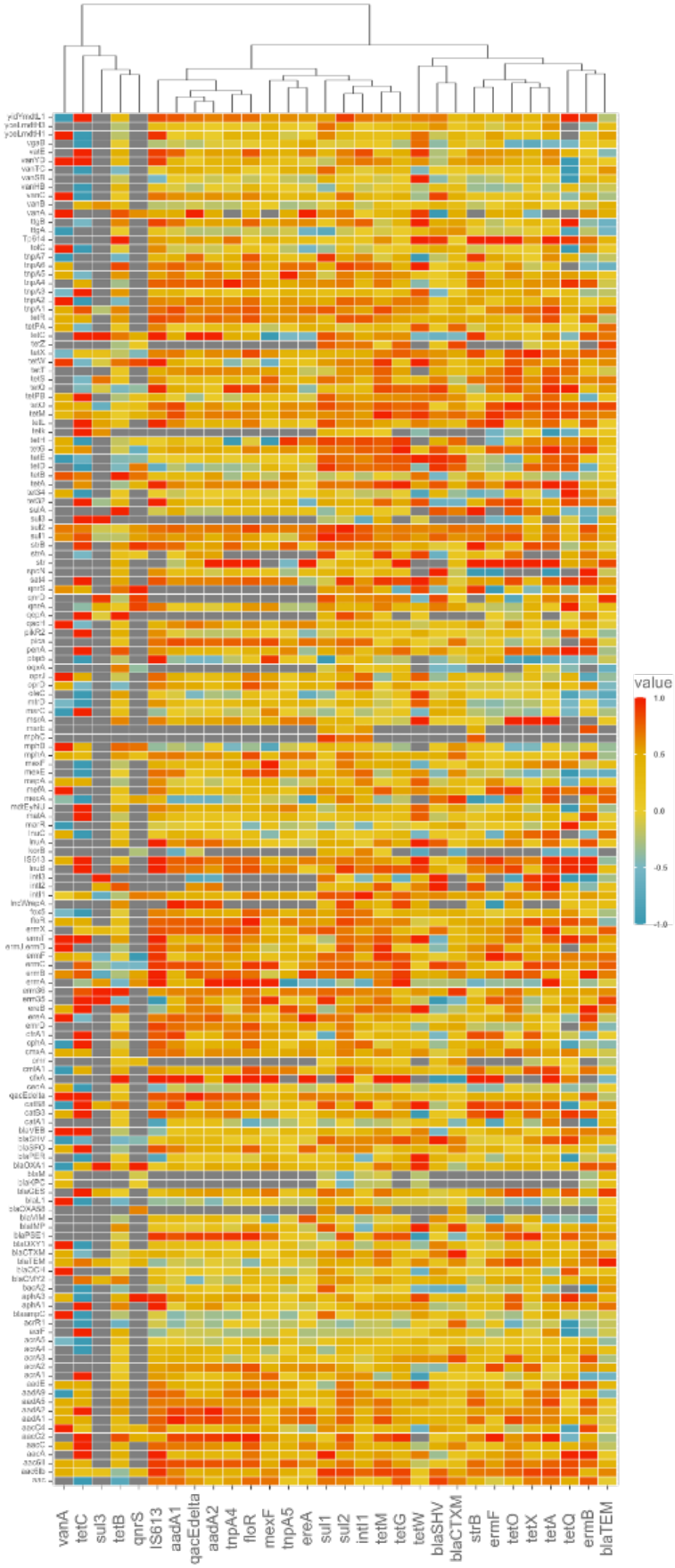



mcr1
Density
N = 154 Bandwidth = 0.5118
mcr1
ARGs/16S rRNA
mphA
Density
N = 108 Bandwidth = 0.4504
mphA
ARGs/16S rRNA
qnrA
Density
N = 142 Bandwidth = 0.6482
qnrA
ARGs/16S rRNA
qnrS
Density
N = 393 Bandwidth = 0.4076
qnrS
ARGs/16S rRNA
strB
Density
N = 141 Bandwidth = 0.4266
strB
ARGs/16S rRNA
sul1
Density
N = 765 Bandwidth = 0.3377
sul1
ARGs/16S rRNA
sul2
Density
N = 590 Bandwidth = 0.3471
sul2
ARGs/16S rRNA
tetA
Density
N = 385 Bandwidth = 0.565
tetA
ARGs/16S rRNA
tetB
Density
N = 156 Bandwidth = 0.3708
tetB
ARGs/16S rRNA
tetG
Density
N = 378 Bandwidth = 0.4081
tetG
ARGs/16S rRNA
tetH
Density
N = 183 Bandwidth = 0.5837
tetH
ARGs/16S rRNA
tetM
Density
N = 427 Bandwidth = 0.466
tetM
ARGs/16S rRNA
air
effluent
feces
food
influent
manure/sludge
other
sediments
soil
water


tetO
Density
N = 353 Bandwidth = 0.7072
tetO
ARGs/16S rRNA
air
effluent
feces
food
influent
manure/sludge
other
sediments
soil
water
tetQ
Density
N = 326 Bandwidth = 0.6812
tetQ
ARGs/16S rRNA
tetW
Density
N = 332 Bandwidth = 0.397
tetW
ARGs/16S rRNA
tetX
Density
N = 276 Bandwidth = 0.4026
tetX
ARGs/16S rRNA
tetC
Density
N = 112 Bandwidth = 0.5229
tetC
ARGs/16S rRNA

850
851

852 **Figure S4. Correlation analysis between the selected ARGs.**



853
854

Figure S5. Relationship between ARG abundances and time. The data points are the abundances of ARGs in samples from each country and the lines are fitted to the maximum abundance values.

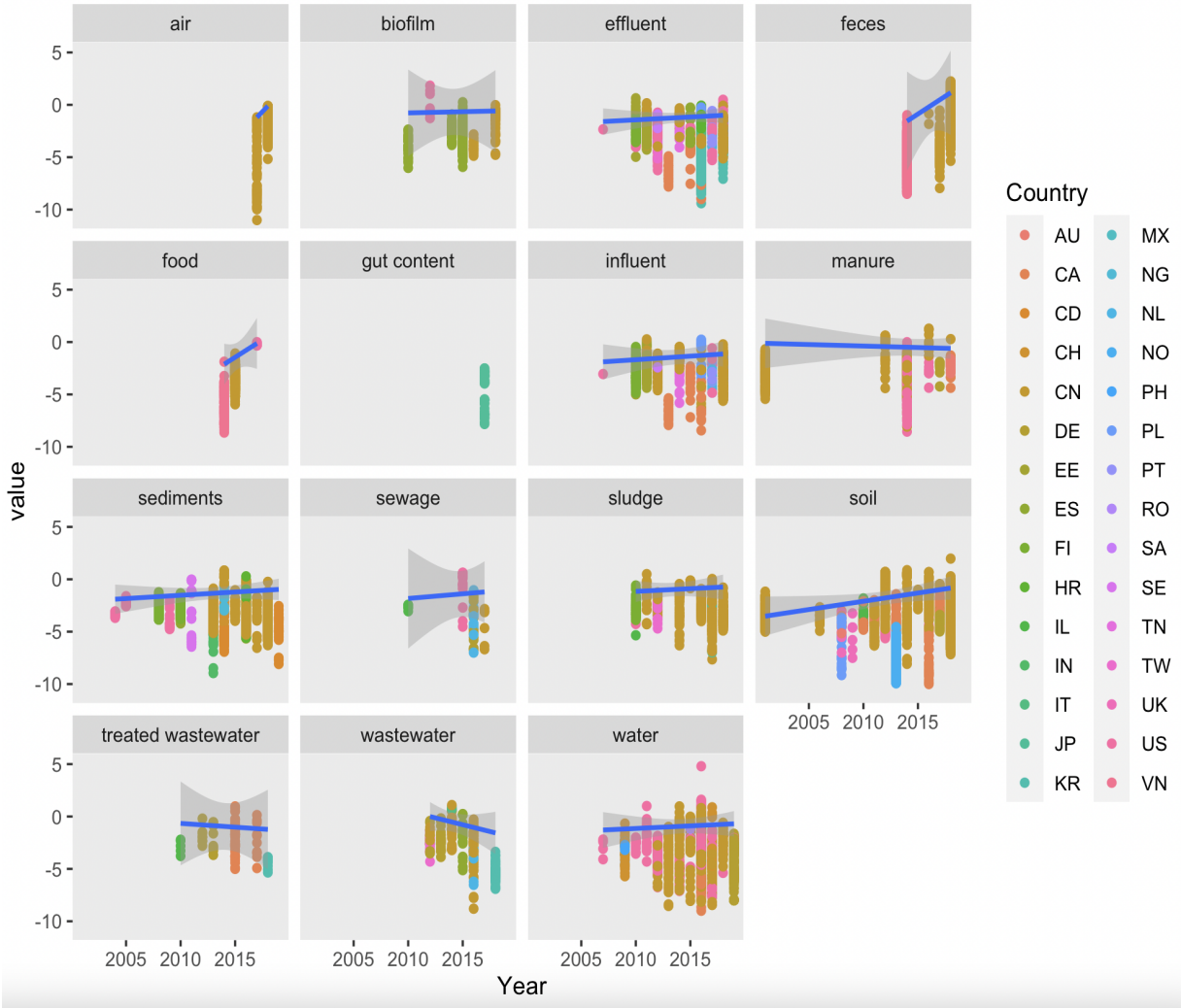
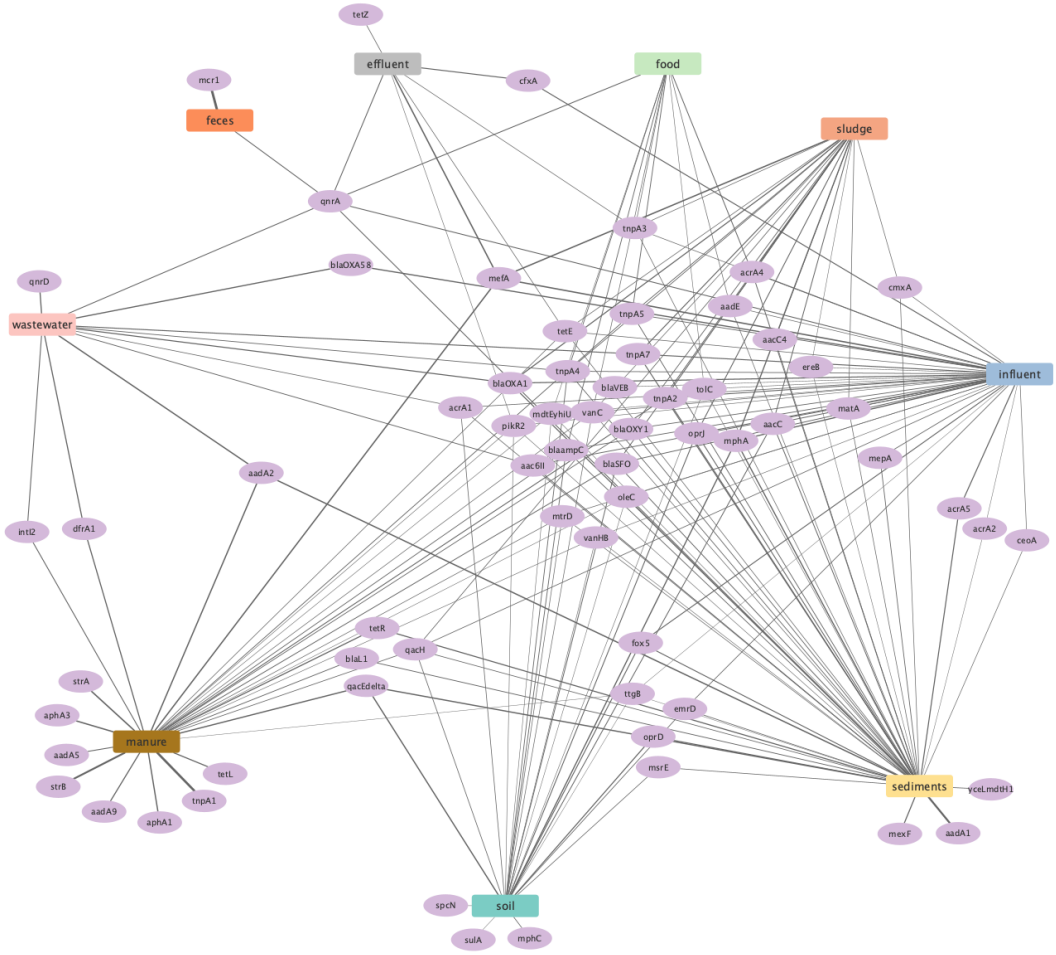


Figure S6. Associations of genes with specific environments. Environments with less than 30 samples have been excluded from the figure.



866
867 **Table S1. Gene classes and the number of unique ARGs they contain**
868

ARG Class	ARGs	Unique ARGs	Reported in studies
Beta lactams		55	78
	blaampC		
	blaAAC		
	blaCMY2		
	blaEC		
	blaNDM		
	blaMOX		
	blaOCH		
	blaPAO		
	blaSM		
	blaROB		
	blaTEM		
	blaCTXM		
	blaOXY1		
	blaSHV2		
	blaPSE1		
	blaAXA1		
	blacepA		
	blacfxA		
	blacphA		
	blaIMP		
	blaVIM		

	blaOXA2		
	blaOXA48		
	blaOXA58		
	blaACT		
	blaL1		
	bla1		
	blaGES		
	blaCMY		
	blaI		
	blaKPC		
	blaLEN		
	blaM		
	blaMIR		
	blaOKP		
	blaOXA1		
	blaOXY2		
	blaPER		
	blaS1		
	blaSFO		
	blaSHV		
	blaSHV34		
	blaTLA		
	blaVEB		
	blaZ		
	carB		

	cepA		
	cfiA		
	cfxA		
	cphA		
	fox5		
	mecA		
	pbp		
	pbp5		
	penA		
Multidrug		46	28
	acrA1		
	acrA2		
	acrA3		
	acrA4		
	acrA5		
	acrB		
	acrF		
	acrR1		
	acrR2		
	adeA		
	adeI		
	qacEdelta		
	ceoA		
	cmeA		
	cmeAF		

	cmle3		
	cmr		
	emrD		
	marR		
	mdtEyhiU		
	mdtg		
	mdth		
	mepA		
	mexA		
	mexB		
	mexE		
	mexF		
	mtrD		
	oprD		
	oprJ		
	oqxA		
	oqxB		
	pmrA		
	qacA		
	qacH		
	rarD		
	sdeB		
	tolC		
	ttgA		
	ttgB		

	yceEmdtG1		
	yceLmdtH1		
	yceLmdtH2		
	yceLmdtH3		
	yidYmdtL1		
	yidYmdtL2		
MLSB		42	75
	cfr		
	ereA		
	ereB		
	erm34		
	erm35		
	erm36		
	ermO		
	ermQ		
	ermA		
	ermB		
	ermC		
	ermF		
	ermJ/ermD		
	ermK		
	ermT		
	ermX		
	ermY		
	lmrA		

	lnuB		
	lmrA		
	lnuF		
	lnuA		
	lnuC		
	matA		
	mdetl1		
	mdtA		
	mefA		
	mefB		
	mphA		
	mphB		
	mphC		
	msrE		
	msrA		
	msrC		
	oleC		
	pikR1		
	pikR2		
	vatA		
	vatB		
	vatC		
	vatD		
	vatE		
	vgaA		

	vgaB		
MGE		40	92
	cIntI1		
	IncNkorA		
	IncNrep		
	IncPoriT		
	IncWrepA		
	intI1		
	intI2		
	intI3		
	IS1133		
	IS1247		
	IS2		
	IS21ISAs29		
	IS256		
	IS26		
	IS3		
	IS6/257		
	IS61		
	IS613		
	ISAba3Acineto		
	ISEfm1Entero		
	ISPps1pseud		
	ISSm2Xanthob		
	korB		

	mobA		
	Tn3		
	TN543		
	Tn916/1545		
	tnpA1		
	tnpA2		
	tnpA3		
	tnpA4		
	tnpA5		
	tnpA6		
	tnpA7		
	tnpA24		
	tnpA25		
	tnpA26		
	tnpA27		
	Tp614		
	traN		
Aminoglycoside		39	44
	aac		
	aac6I		
	aac6Ib		
	aac6II		
	aac6Iy		
	aac6iz		
	aac2ic		

	aac3ia		
	aac3iv		
	aac3vi		
	aac6Ia		
	aacA		
	aacC		
	aacC1		
	aacC2		
	aacC4		
	aadA1		
	aadA2		
	aadA5		
	aadA9		
	aadD		
	aadE		
	acrD		
	ant2ia		
	ant3ia		
	aph		
	aph2Id		
	aph3ia		
	aph3III		
	aph3VIa		
	aph4Ib		
	aph6Ia		

	aphA1		
	aphA3		
	sat4		
	spcN		
	str		
	strA		
	strB		
Tetracycline		31	116
	tet32		
	tet34		
	tet35		
	tet36		
	tet37		
	tet38		
	tetA		
	tetB		
	tetD		
	tetE		
	tetG		
	tetH		
	tetJ		
	tetk		
	tetL		
	tetM		
	tetO		

	tetPB		
	tetQ		
	tetS		
	tetT		
	tetU		
	tetV		
	tetW		
	tetX		
	tetZ		
	tet39		
	tet44		
	tetC		
	tetPA		
	tetR		
Vancomycin		21	18
	vanA		
	vanB		
	vanC		
	vanG		
	vanHB		
	vanHD		
	vanRA		
	vanRB		
	vanRC		
	vanRC4		

	vanSB		
	vanSC		
	vanTC		
	vanTE		
	vanTG		
	vanWB		
	vanWG		
	vanXA		
	vanXD		
	vanYB		
	vanYD		
Biocide/Metal		16	16
	arsA		
	cadA		
	chrA		
	copA		
	crcA		
	czcA		
	fabK		
	gadAB		
	merA		
	nimE		
	nisB		
	pcoA		
	pncA		

	qacF		
	rcnA		
	tcrB		
Phenicols		10	16
	catA1		
	catB3		
	catB8		
	catQ		
	cmlA1		
	cmlA5		
	cmlV		
	cmxA		
	fexB		
	floR		
Trimethoprim		8	15
	dfrA1		
	dfrA12		
	dfrA13		
	dfrA14		
	dfrA17		
	dfrA22		
	dfrC		
	folA		
Quinolones		5	37

	qepA		
	qnrA		
	qnrB		
	qnrD		
	qnrS		
Sulfonamide		4	112
	sul1		
	sul2		
	sul3		
	sulA		
Bacitracin		2	8
	bacA1		
	bacA2		
Colistin		1	5
	mcr1		
Other		7	5
	aerA3		
	EAE5855		
	gyr2		
	Pe		
	pica		
	trbC		
	uidA		

869
870

871 **Table S2. Abundance distribution for reported ARGs**
872

Gene	Number of observations	Min	1Q	Median	Mean	3Q *	Max
aac	26	-4,9519212	-4,4757121	-4,2218488	-3,9104352	-3,8467018	-0,4484305
aac6I	13	-5,1149177	-4,457575	-4,3129996	-4,2296892	-3,7864876	-3,5455649
aac6Ib	112	-9,4	-3,7017708	-3,2348348	-3,5147957	-2,4674718	0,832994
aac6II	60	-9,34838	-4,4838731	-3,6406975	-3,9145249	-3,0994286	-1,4496602
aac6Iy	6	-5,5415792	-5,3246863	-5,2065132	-5,0372733	-4,9859138	-3,992576
aac2ic	8	-9,6499685	-8,4514287	-8,3158796	-8,3424101	-8,1023244	-7,4584276
aac3ia	7	-7,9318141	-7,7356733	-4,3870968	-4,6428478	-1,6451613	-1,4193548
aac3iv	6	-8,9194889	-8,1117715	-7,4712342	-7,6428411	-7,2043659	-6,5543958
aac3vi	8	-7,3165287	-6,1712173	-5,6683086	-5,6983329	-5,3215454	-3,9389987
aac6Ia	6	-8,7572714	-8,4378682	-8,0273174	-7,9450309	-7,4791277	-6,9871628
aacA	26	-4,3979487	-4,1868759	-3,7039003	-3,5162445	-3,1572828	-1,6575773
aacC	76	-5,9264869	-3,3427102	-2,4915878	-2,8260093	-1,9787251	-1,4243375
aacC1	8	-4,6478175	-4,4836352	-3,5679758	-3,7868803	-3,2467236	-2,9976637
aacC2	32	-5,126659	-4,0413708	-3,633275	-3,4912729	-2,9089268	-1,4367186
aacC4	42	-4,8354651	-3,8714078	-3,4859938	-3,2875111	-2,8651603	-0,4516129
aadA1	157	-8,4920128	-3,4800069	-2,8413476	-2,9429659	-2,0930233	-0,1503859
aadA2	109	-5,2146717	-3,4669575	-2,7242572	-2,7948976	-2,0982803	-0,8927934
aadA5	80	-8,2572749	-3,8354072	-3,3806947	-3,6959273	-2,7783093	-1,4546929
aadA9	52	-8,1272612	-5,4339749	-4,1509505	-4,5176717	-3,6642403	-1,9577465
aadD	17	-4,9144635	-4,242761	-3,7996539	-3,4780319	-2,4904596	-1,1873199
aadE	51	-4,853872	-3,9393305	-3,6289321	-3,5091075	-3,0525579	-1,6925396
acrA1	62	-9,7983176	-4,5694098	-3,6407291	-4,0731819	-3,2571287	-2,3896669
acrA2	36	-4,7581743	-3,8011755	-3,3906826	-3,3043091	-2,8582465	-1,2423775
acrA3	18	-5,5759434	-4,6322954	-4,2979081	-4,3284233	-3,9331025	-3,3575355
acrA4	56	-4,9871628	-3,42507	-3,13305	-3,1807451	-2,8515393	-1,3147428
acrA5	55	-4,8893869	-3,5638365	-3,1713401	-3,1925855	-2,8674122	-1,5985566
acrB	30	-9,377786	-5,0522255	-3,3723788	-4,0757612	-2,4374377	-0,14
acrD	2	-7,6194762	-7,5538031	-7,4881301	-7,4881301	-7,422457	-7,3567839
acrF	31	-8,6842965	-5,3889386	-4,1549196	-4,6841829	-3,2686842	-2,5925991
acrR1	45	-8,7878124	-4,6595559	-3,756962	-4,3560271	-3,4469199	-2,5100415
acrR2	22	-5,1726813	-4,320921	-3,3575616	-3,0495286	-1,5801535	-0,5542748
adeA	17	-8,4282912	-7,6126117	-4,4881166	-5,3356966	-3,6539834	-3,1357157
aerA3	2	-4,2564679	-4,0542826	-3,8520973	-3,8520973	-3,649912	-3,4477266
blaampC	92	-11	-4,0899574	-3,5369584	-3,7777532	-3,2555934	-0,6667569
ant2ia	5	-8,1396338	-7,8761484	-7,6252517	-7,4999105	-7,5451551	-6,3133637
ant3ia	6	-9,347983	-8,4464652	-8,1645803	-8,0269659	-7,8280046	-6,2284125
aph	5	-4,2218488	-3,8944898	-3,8461851	-3,015606	-1,7264574	-1,389049
aph2Id	5	-4,5015769	-2,6634215	-2,5751184	-2,8457754	-2,3498868	-2,1388734
aph6Ia	6	-8,1366771	-7,5961104	-5,3687942	-5,2353065	-2,77002	-2,2952624

aphA1	54	-5,8450704	-4,3147674	-3,494115	-3,5306465	-2,8955238	-1,2580645
aphA3	41	-5,7356828	-4,2218488	-3,3943662	-3,4935269	-3,056338	-0,3043478
bacA1	20	-7,8612914	-6,6266514	-4,4431466	-4,9368967	-3,6814638	-2,6289321
bacA2	28	-4,9691139	-4,13652	-3,4193046	-3,6168816	-3,2546558	-2,5050828
blaAAC	14	-9,2533658	-7,6850985	-4,4227994	-5,2930264	-4,275569	-1
blaCMY2	56	-8,5655485	-4,0731693	-3,6705772	-3,8480256	-3,2590396	-0,2798407
blaEC	7	-9,8694587	-8,5215603	-7,4412914	-7,9287559	-7,4050893	-6,3372422
blaNDM	47	-7,8613992	-6,5234999	-6,1733434	-5,8729964	-5,2381034	-3,3254902
blaMOX	19	-8,8134291	-4,3412505	-3,6845546	-3,8793275	-2,9517765	-2,5365646
blaOCH	24	-7,9439515	-4,6432774	-4,1915267	-4,5798147	-3,8035999	-3,3225839
blaPAO	8	-9,3338913	-7,8364708	-4,8828076	-5,6944753	-4,1300407	-2,5434821
blaSM	5	-6,9788179	-6,8961963	-6,8612914	-6,7511405	-6,7644716	-6,2549253
blaROB	21	-8,5214335	-7,4116928	-4,6005435	-5,5607359	-4,279847	-3,1817864
blaTEM	370	-9,8	-4,4157895	-3,4641077	-3,6297081	-2,6271632	1,97175361
blaCTXM	324	-9	-5,419205	-4,0809511	-4,3740689	-3,3140691	0,642455
blaOXY1	64	-7,9244534	-3,8213825	-3,2255231	-3,4123879	-2,6392158	-2,1257234
blaSHV2	16	-8,3458235	-6,8532003	-4,1709511	-4,9708162	-3,4322738	-3,1299724
blaPSE1	26	-9,7491569	-4,6517062	-3,9635711	-4,5068836	-3,4256515	-1,7
blaAXA1	2	-8,7585898	-8,6271005	-8,4956113	-8,4956113	-8,364122	-8,2326327
blacepA	24	-7,6675615	-5,375	-5,265625	-5,0293392	-4,8359375	-3,1870666
blacfxA	17	-8,2418818	-6,7423214	-3,9199687	-4,6323485	-2,9520865	-1,3710988
blacphA	21	-8,3936186	-6,391474	-3,7836514	-4,7724305	-3,4392422	-2,139662
blaIMP	32	-8,7721133	-4,1375602	-3,7746479	-3,9213569	-3,2409997	-2,5166415
blaVIM	21	-9,9136417	-6,872467	-4,761875	-5,4489683	-4,2856724	-2,853872
blaOXA2	14	-3,853872	-3,2468509	-2,6477416	-2,7427816	-2,3933399	-1,46
blaOXA48	18	-8,1419661	-6,8861805	-5,2053733	-5,6661858	-4,6885397	-2,8860566
blaOXA58	63	-8	-5,3794212	-3,0470588	-3,8263565	-2,401118	-0,8972332
blaL1	26	-4,9691139	-4,1896075	-3,8405332	-3,8223982	-3,5196756	-2,4629275
bla1	7	-4,3385827	-3,7626595	-3,5272436	-3,5523093	-3,1762427	-3,1225348
blaGES	52	-8,1349626	-4,4067049	-3,3018831	-3,5838719	-2,4912531	-1,1977569
blaCMY	18	-4,7429794	-4,2236942	-3,9066702	-3,8436073	-3,4414847	-2,7805278
blaI	9	-8,7996539	-8,3223935	-7,6289321	-7,5297612	-7,1992829	-4,3329287
blaKPC	35	-6,79588	-4,571941	-4,3451627	-4,2007091	-3,8617188	-1,4641784
blaM	24	-4,2422226	-3,4727048	-3,2069838	-2,7653714	-2,6125222	0,24891454
blaOKP	6	-5,2415834	-4,0908536	-3,8149419	-4,0017292	-3,7158567	-3,2732557
blaOXA1	99	-8,084507	-4	-3,3321838	-3,4596452	-2,6746968	-0,8552742
blaPER	24	-4,9767632	-4,3351809	-4,1740892	-3,9742154	-3,5960247	-3,1234935
blaS1	10	-3,5872892	-3,3580467	-3,1647432	-3,1425506	-2,8850679	-2,7453715
blaSFO	76	-6,3539482	-4,0139612	-3,154927	-3,5084731	-2,7850277	-1,6646986
blaSHV	128	-7,9	-5,2124389	-3,7363564	-4,0173508	-3,021704	0,05882353
blaSHV34	8	-2,6536473	-2,5598003	-2,429633	-2,1699697	-2,136826	-0,5421181
blaTLA	10	-4,6514926	-4,1030371	-3,5590519	-3,7143376	-3,3807557	-3,0352346
blaVEB	27	-4,7746972	-3,7551278	-3,3266218	-3,0462076	-2,2656466	-0,9753456

blaZ	8	-5,5818133	-5,2320882	-5,0186688	-5,0601195	-4,8226441	-4,6738916
cadA	2	-7,6194762	-7,5538031	-7,4881301	-7,4881301	-7,422457	-7,3567839
carBR	7	-9,1294125	-8,8951746	-8,5467512	-8,4301032	-8,3431604	-6,8578887
catA1	19	-4,5957915	-4,2092942	-3,9691139	-3,9051943	-3,6972801	-2,682459
catB3	39	-9,5724772	-3,8905281	-3,5148865	-3,9456811	-2,9506131	-1,338237
catB8	26	-8,7121983	-3,8974557	-3,4971436	-3,8394211	-3,0643817	-2,464276
qacEdelta	120	-8,844664	-3,3187531	-2,4819809	-2,7126507	-1,7741203	-0,0140845
ceoA	60	-8,2395775	-3,9456964	-3,6016247	-3,9255556	-3,2327188	-2,7734948
cepA	9	-4,5666736	-4	-3,7528496	-3,805982	-3,5168945	-3,3898928
cfiA	12	-8,9443978	-5,7524547	-4,5586412	-5,3652273	-4,0057294	-3,6989743
cfr	19	-5,045256	-1,6	-0,0762763	-1,1162645	0,05913301	0,38512222
cfxA	39	-5,5147436	-3,8264397	-3,5649401	-3,4376749	-3,0915908	-0,8
chrA	2	-7,1758794	-6,9949507	-6,814022	-6,814022	-6,6330932	-6,4521645
cIntI1	7	-1,7706635	-1,4015458	-1,2364901	-1,1948657	-0,9866868	-0,5804411
cmeA	2	-5,4793796	-5,0273253	-4,5752709	-4,5752709	-4,1232166	-3,6711623
cmeAF	1	-3,8632794	-3,8632794	-3,8632794	-3,8632794	-3,8632794	-3,8632794
cmle3	7	-9,4958677	-7,8539813	-6,6345122	-7,2505934	-6,559965	-5,7958817
cmlA1	87	-8,4736672	-3,9396562	-2,9912759	-2,9963404	-1,9332355	0,38512222
cmr	15	-6,7	-6,1	-5,3	-4,9172566	-4,2	-1,7
cmxA	53	-4,883213	-4,3564978	-3,7759852	-3,7303478	-3,3593656	-1,394695
copA	1	-0,25451	-0,25451	-0,25451	-0,25451	-0,25451	-0,25451
cphA	68	-4,928582	-3,6220746	-3,2670164	-3,2503907	-2,8053226	-1,1818139
crcA	1	-0,1414851	-0,1414851	-0,1414851	-0,1414851	-0,1414851	-0,1414851
dfrA1	74	-9,3488328	-4,2182613	-3,3287419	-3,63422	-2,4831123	-1,4221643
dfrA12	18	-4,7281584	-4,3055341	-4,04875	-3,9910516	-3,772736	-2,9535735
dfrA14	8	-3	-2,395	-2,01	-1,9675	-1,6	-0,98
dfrA22	9	-5,1545455	-5	-3,8534483	-4,0760152	-3,2829651	-3,173913
emrD	57	-9,2741984	-4,7713457	-3,3129996	-3,7671441	-2,568688	-1,7287489
ereA	93	-5,7536257	-3,6298389	-3,1898844	-2,8672118	-2	1
ereB	37	-8,8477117	-4,4841262	-3,8356471	-4,2853293	-3,3791447	-2,443281
erm34	13	-8,7987667	-7,6695862	-4,5373519	-5,4458531	-4,1769623	-2,8896839
erm35	21	-5,5425699	-4,4019682	-3,8632794	-3,9117773	-3,3212662	-2,5269969
erm36	52	-8,2211255	-5,1621572	-3,8102132	-4,2993815	-3,2722702	-1,565947
ermA	29	-8,799559	-4,543634	-3,7659168	-4,1245782	-2,8454918	-2,2745551
ermB	329	-9	-3,4977625	-2,6594954	-2,8298469	-1,8526316	0,96214585
ermC	69	-8	-4,7372665	-3,0988142	-3,7588676	-2,4953826	-1,2235294
ermF	132	-7,7695518	-3,3183742	-2,3476248	-2,4713145	-1,348239	-0,0462403
ermJ.ermD	13	-3,9685915	-3,4720877	-3,2616159	-3,3138558	-3,1643943	-2,7264254
ermK	11	-9,8865665	-6,0348983	-4,4682944	-5,2750132	-4,0539917	-2,9943914
ermT	34	-4,8927934	-4,2388007	-3,9867639	-3,8542724	-3,58861	-1,9691138
ermX	67	-7,2684112	-4,2037123	-3,3592194	-3,4562063	-2,4078085	-1,5799236
ermY	8	-4,3467875	-4,1446568	-3,9642989	-4,011612	-3,8691077	-3,7695518
fabK	5	-9,2641377	-8,5128616	-8,3882767	-7,0378448	-4,6259991	-4,3979487

fexB	14	-3,8249515	-0,4682028	-0,2982323	-0,9201291	-0,1521647	0,62546438
floR	100	-6,1943918	-3,9093381	-3,2617171	-2,8596675	-2,0702723	0,6213178
folA	10	-5,61	-4,9325	-4,2723214	-4,3894643	-3,7975	-3,36
fox5	76	-9,3747383	-4,7305809	-3,1970383	-3,692453	-2,7072836	-1,7189666
gadAB	5	-4,7	-4,5	-3,8	-4,06	-3,7	-3,6
gyr2	5	-4,4259687	-3,6989743	-3,2218488	-3,0091532	-2,6989743	-1
lmrA	4	-4,3233639	-4,2472275	-3,9447051	-3,9515585	-3,6490361	-3,5934598
IncWrepA	17	-6,5211268	-4,8136364	-4,4347826	-4,452097	-3,5016894	-2,7935261
intI1	558	-9	-3,3	-2,3186645	-2,4605839	-1,3984615	1,6
intI2	30	-6,8831169	-4,796847	-3,9961965	-3,736675	-2,735369	-0,5
intI3	17	-6,9264693	-5,1445934	-3,4588161	-3,6255432	-2,6338028	-0,6901408
lnuB	41	-8,3979487	-3,7996539	-3,3494623	-3,4781035	-2,8223531	-1,9439515
IS613	44	-5,7117124	-4,3112874	-3,5553119	-3,3845149	-2,4690811	-0,7144427
korB	15	-4,24	-3,545	-3,29	-3,1670795	-2,8631237	-1,8441558
lmrA	1	-8,2313619	-8,2313619	-8,2313619	-8,2313619	-8,2313619	-8,2313619
lnuA	26	-6	-4,3776291	-3,8476912	-3,9238597	-3,5159762	-2,2783521
lnuC	24	-8,263635	-6,4175808	-4,6070531	-5,0548177	-4,0725676	-2,4221643
m1	33	-9,3364361	-4,6736641	-2,977268	-3,7881121	-2,5697474	-0,5493375
marR	24	-8,4377714	-5,8865797	-4,4831167	-5,1111517	-4,0721631	-2,9318141
matA	63	-8,9661145	-3,7730543	-3,4424928	-3,6549946	-2,6801942	-1,6185579
mdetI1	8	-4,8354651	-4,0866969	-3,8986526	-3,9785206	-3,7261774	-3,4958677
mdtA	9	-9,7695518	-6,7114692	-4,8447377	-5,6509908	-3,8827287	-3,1463179
mdtEyhiU	56	-7,4596753	-4,4101237	-3,7267907	-4,0854895	-3,2921318	-2,4638156
mecA	31	-7,920904	-6,5329057	-4,5228787	-4,7377975	-3,3734982	-0,2218488
mefA	120	-9,9251478	-3,7135486	-2,748603	-2,9429358	-1,5392058	-0,254625
mepA	63	-7,8865665	-3,8365972	-3,5199931	-3,7904286	-3,2034977	-1,4254439
merA	2	-7,2727273	-7,2485153	-7,2243033	-7,2243033	-7,2000914	-7,1758794
mexA	17	-8,7262964	-7,455932	-5,1824591	-5,470074	-3,7986288	-1,3
mexB	2	-6,6934673	-6,6672652	-6,6410631	-6,6410631	-6,614861	-6,5886589
mexE	38	-8,2629615	-5,5317143	-4,367695	-4,9469165	-3,9138445	-3,4596753
mexF	84	-6,1438756	-2,8493666	-2,220371	-2,5306468	-1,7572611	-1,2328441
mcr1	154	-8,4945434	-6,81646	-5,6887561	-5,6048708	-4,7296834	-2,0627351
mphA	108	-6,9657831	-4,3380631	-2,868606	-3,3589138	-2,4028707	-1,6193818
mphB	28	-8,9194889	-4,9600395	-4,4399778	-4,90916	-3,7739709	-2,5196337
mphC	32	-9,2153828	-4,4817915	-4,1935496	-4,2751543	-3,5788722	-2,4526788
msrE	44	-4,7859922	-3,4839494	-3,0471681	-2,9622431	-2,5808547	-0,1
msrA	19	-4,8624149	-4,5154085	-4,1549196	-4,0907898	-3,6648042	-3,3396151
msrC	30	-8,7746972	-5,6977763	-4,2770997	-4,8745868	-3,8229581	-1,2
mtrD	53	-9,7983176	-5,2757241	-3,7825723	-4,4779232	-3,3724661	-2,2848479
nimE	8	-4,8687542	-4,4689075	-4,3103563	-4,2391671	-4,0574653	-3,5767541
nisB	10	-5,9782568	-4,6195616	-4,2391418	-4,3495379	-3,7407879	-3,5342916
oleC	41	-7,6733467	-4,2424731	-3,329453	-3,9782998	-3,1356379	-2,5316527
oprD	78	-8,3272758	-4,9150927	-3,2420128	-3,7458889	-2,53695	-1,1296789

oprJ	85	-9,7983176	-3,6838786	-2,826258	-3,3195944	-2,3838499	-1,6126117
oqxA	12	-8,5	-8,1	-7,7	-7,6666667	-7,475	-6,4
oqxB	15	-5,7865613	-4,4782204	-3,971831	-3,6008572	-2,9599385	0,197158
pbp	10	-5,3120301	-4,09698	-2,4304932	-3,1089294	-2,3268078	-2,1295969
pbp5	12	-4,8094445	-4,2508616	-3,9366504	-4,0107935	-3,6955019	-3,4762535
pcoA	1	-0,2848365	-0,2848365	-0,2848365	-0,2848365	-0,2848365	-0,2848365
Pe	10	-4,1837541	-3,5878742	-3,4025898	-3,4220243	-3,0575072	-2,9356766
penA	38	-7,7986288	-6,4075	-4,4403721	-4,8199716	-3,6153585	-2,7746907
pica	23	-5,7431367	-4,9770035	-4,4249577	-3,9392326	-2,7211733	-2,3728929
pikR1	13	-9,3526979	-7,8865665	-7,733349	-6,35675	-4	-3,251812
pikR2	48	-7,6615436	-3,9482714	-3,5450468	-3,9798384	-3,2520815	-2,8416375
pmrA	9	-5,4727296	-3,8312798	-3,7632196	-3,9266196	-3,5852665	-3,1374944
pncA	49	-7,756962	-3,8144471	-3,2567356	-3,6334207	-2,8514444	-2,1778319
putitive.multidrug	11	-4,2411998	-3,9001858	-3,4436975	-3,4870363	-3,0346401	-2,5844252
qacA	7	-8,6798537	-6,8291018	-4,4042612	-5,5860711	-4,3309733	-3,6982323
qacF	2	-6,5367965	-6,3950597	-6,2533229	-6,2533229	-6,1115861	-5,9698492
qacH	69	-8,2321238	-4,5256628	-3,2422246	-3,8369603	-2,563812	-1,6454398
qepA	32	-5,9389987	-3,0297479	-2,5613262	-2,5493383	-1,9930486	0,197158
qnrA	142	-8,3687529	-6,845024	-4,9638346	-5,0229518	-3,4740046	-1,1764706
qnrB	26	-5,6526316	-4,3501921	-3,4238688	-3,5862685	-2,8243778	-1,9705882
qnrD	41	-7,7	-6,4	-4,7835655	-4,763266	-2,9963613	-1,8878924
qnrS	393	-9,1	-5,0435554	-4,1	-4,1631106	-3,038961	0,47058824
rarD	22	-8,7522673	-4,0337549	-3,6006563	-4,1338242	-3,1947056	-2,5725607
rcnA	2	-7,4891775	-7,4560791	-7,4229807	-7,4229807	-7,3898823	-7,3567839
sat4	30	-4,7996539	-4,2991914	-3,641241	-3,6646016	-2,9332013	-2,1237822
sdeB	17	-7,8134291	-3,9389987	-3,7269987	-4,1828991	-3,4328898	-3,1415634
spcN	25	-8,9665762	-7,26984	-5,2695474	-5,49099	-3,9685915	-3,1242878
str	18	-4,4271284	-3,9057186	-3,3187413	-3,2537543	-2,6871732	-2
strA	75	-9,1446281	-4,2287413	-3,7260682	-3,4447449	-2,0908467	-0,3625103
strB	141	-8,6193469	-3,7958817	-2,8124793	-3,085025	-2,0869565	-0,5647059
sul1	765	-8,9	-3,1249387	-2,1788618	-2,218761	-1,227451	4,8
sul2	590	-9,1146388	-3,520734	-2,6721964	-2,6869709	-1,6696442	1,8260748
sul3	39	-5,5855379	-3,3331239	-2,8111366	-2,7216294	-2,0587677	0,4557973
sulA	34	-8,8398976	-6,4318696	-5,27	-5,7033514	-4,753757	-3,2328441
tet32	60	-4,7524423	-3,3637245	-2,8587124	-3,0098629	-2,5860722	-1,8631433
tet34	44	-9,1649439	-4,3320343	-3,8832	-4,272518	-3,5286535	-1,2656578
tet35	1	-4,5883829	-4,5883829	-4,5883829	-4,5883829	-4,5883829	-4,5883829
tet36	18	-4,8469236	-4,0949094	-3,8865665	-3,9235627	-3,6828237	-3,2141929
tet37	13	-8,7619539	-7,9546772	-4,457575	-5,1143805	-2,7271419	-1,3281819
tet38	1	-4,4259687	-4,4259687	-4,4259687	-4,4259687	-4,4259687	-4,4259687
tetA	385	-8,9578198	-3,4343434	-1,7687854	-2,0825628	-0,374125	1,85847704
tetB	156	-8,7958817	-4,0110132	-3,4067165	-3,3829141	-2,495229	0,48482513

tetD	53	-8,8153857	-6,5	-4,6259991	-5,084205	-3,8167316	-1,8436364
tetE	72	-8,7798919	-5,5018888	-4,0798467	-4,3997362	-2,9836115	-1,161435
tetG	378	-8,3936186	-3,1787495	-2,3714722	-2,2879638	-1,1872499	2,2008505
tetH	183	-9,8275433	-3,2227397	-2,024844	-2,2259687	-0,7593561	1,03981055
tetJ	4	-5,0042759	-4,2351824	-3,9590698	-4,1992026	-3,92309	-3,8743951
tetk	19	-8,5345977	-3,2017457	-2,7226473	-3,238397	-2,6025475	-1,7383996
tetL	86	-9,5898576	-4,0605076	-3,0799566	-3,4240705	-2,3787083	-0,1526718
tetM	427	-9,1191865	-4,494399	-3,2	-3,3500243	-2,164459	1,09221902
tetO	353	-10	-3,5271433	-2,2344892	-1,918661	0,33041377	2,15014216
tetPB	60	-8,7258422	-4,7	-3,8531806	-4,2325497	-3,3592	-2,671624
tetQ	326	-8,0712121	-2,8427497	-1,4428991	-1,5486061	0,58103895	2,23502316
tetS	51	-4,9643519	-4,1353584	-3,3492696	-3,3915836	-2,8998718	-1,0906948
tetT	32	-5,116127	-4,435937	-3,8360699	-3,570502	-3,2013359	-0,8396947
tetU	4	-8,6326279	-7,7772651	-7,480988	-7,1504514	-6,8541743	-5,0072016
tetV	5	-4,8354651	-4,5666736	-4,279847	-4,2100269	-4,139662	-3,2284867
tetW	332	-8,1747368	-3,9514789	-2,8229127	-3,0437981	-2,063991	0,67759664
tetX	276	-9,6612946	-3,764144	-2,5372043	-2,8407957	-1,9196489	0,70141674
tetZ	24	-5,9	-5,4	-4,3	-4,0233388	-2,8344077	-0,9281875
tet39	7	-3,99	-3,015	-2,19	-2,0971429	-1,375	0,28
tetC	112	-5,48	-3,6292797	-2,7723529	-2,560247	-1,6196716	0,63992522
tetPA	27	-4,2999069	-3,8704301	-3,5697823	-3,514727	-3,2457368	-2,5844252
tetR	90	-7,4692391	-4,5936188	-3,2876524	-3,7874239	-2,9514343	-1,6757175
tnpA1	95	-6,6056338	-4,3687145	-3,2319752	-3,0276395	-2,2577933	-0,1135376
tnpA2	69	-5,9685294	-3,7695518	-3,3391345	-3,2671155	-2,5624647	-1,1
tnpA3	63	-5,8459139	-3,6570035	-3,3122538	-3,0986786	-2,4440694	-0,465575
tnpA4	81	-4,6124482	-3,3788237	-2,552842	-2,5805286	-1,7466162	-0,7717788
tnpA5	72	-5,3982669	-3,6565714	-2,8824055	-2,7650906	-1,7945112	-0,7697893
tnpA6	34	-5,6221894	-3,974026	-3,0085936	-3,1572054	-2,6565282	-1,1691273
tnpA7	65	-5,1379625	-3,5354436	-2,7475389	-2,912904	-2,2718196	-1,5397139
tolC	54	-8,1396338	-4,1796297	-3,7339183	-4,0883324	-3,3678387	-2,5974748
Tp614	43	-6,1852119	-4,382708	-3,4122212	-3,4033446	-2,3296932	-0,8700133
ttgA	32	-8,733349	-4,8931103	-4,1810699	-4,8715991	-3,9296981	-3,1366771
ttgB	38	-8,6499685	-5,6875218	-4,411502	-4,9822019	-3,9419088	-2,9756554
uidA	5	-4,9	-4,6	-3,6	-3,66	-3,6	-1,6
vanA	26	-4,9691139	-3,2970107	-2,3896854	-2,3162074	-1,2550536	0,26245109
vanB	64	-7,8694587	-3,9544912	-3,5363188	-3,9753566	-3,1820009	-2,1545424
vanC	67	-7,8239875	-3,6301961	-2,8413215	-3,327126	-2,6240706	-2,1667916
vanG	3	-4,4881166	-4,3379915	-4,1878664	-4,218267	-4,0833422	-3,9788179
vanHB	42	-7,7544873	-4,1762167	-3,6889915	-4,1342044	-3,3535687	-2,1954289
vanHD	5	-8,4221643	-8,2564924	-7,9151498	-6,6300255	-4,3129996	-4,2433216
vanRA	9	-4,6259991	-4,3467875	-4,1493538	-4,062933	-3,9546772	-2,5668187
vanRB	9	-8,3142583	-7,7399286	-5,4107448	-6,0590979	-4,457575	-4,116983
vanRC	2	-5,5868778	-5,5104379	-5,433998	-5,433998	-5,3575582	-5,2811183

vanSB	25	-8,5917635	-4,6345122	-4,1943274	-4,8680501	-3,9625735	-2,3493354
vanSC	12	-9,5421181	-8,3015527	-4,77639	-5,9996954	-4,1147686	-3,65267
vanTC	24	-8,4156688	-7,3559786	-4,0727908	-5,233389	-3,8530716	-3,4528488
vanTE	2	-8,6736641	-7,4949164	-6,3161686	-6,3161686	-5,1374209	-3,9586731
vanTG	4	-8,6439741	-5,7072625	-4,6306355	-5,5545612	-4,4779342	-4,3129996
vanWB	4	-9,6675615	-5,8418956	-4,2529976	-5,3901932	-3,8012952	-3,3872161
vanWG	18	-4,9674841	-4,1820246	-3,7359799	-3,4458515	-2,7	-1,4
vanXA	2	-9,2276783	-9,0570475	-8,8864167	-8,8864167	-8,7157859	-8,5451551
vanXD	12	-5,6651271	-5,3954207	-5,0292321	-4,5799358	-3,6340027	-2,7159564
vanYB	10	-9,8413476	-9,0790905	-7,2620702	-6,9918029	-4,7254346	-4,1007987
vanYD	30	-8,4377714	-5,4297866	-4,0068331	-4,8102204	-3,676235	-3,2134935
vatB	2	-4,455932	-3,9143668	-3,3728017	-3,3728017	-2,8312366	-2,2896714
vatC	6	-4,6259991	-4,5489226	-4,3896479	-4,4143419	-4,275569	-4,2411998
vatD	2	-3,8312798	-3,6693995	-3,5075193	-3,5075193	-3,345639	-3,1837588
vatE	39	-9,7258422	-5,7627968	-4,279847	-4,7755884	-3,3557418	-1,3129996
vgaA	13	-9,9691139	-7,6989743	-4,6478175	-5,8281677	-4,1549196	-3,3952341
vgaB	23	-7,9318141	-6,3450359	-4,3129996	-4,8974519	-3,8535059	-2,6363882
yceEmdtG1	14	-4,9691139	-3,8268428	-3,7181892	-3,5742862	-2,9839538	-2,5474412
yceLmdtH1	38	-8,4449555	-4,1769623	-3,7194468	-4,3053937	-3,413565	-2,6516951
yceLmdtH2	17	-4,935294	-4,5666736	-4,123327	-3,9623303	-3,6458916	-2,649545
yceLmdtH3	24	-4,4881166	-4,2220775	-3,7596053	-3,7089243	-3,3075991	-2,6615435
yidYmdtL1	37	-8,9829667	-6,3891858	-4,1995384	-4,9366458	-3,756962	-2,7694444
yidYmdtL2	10	-4,4723799	-4,3190911	-3,7537676	-3,6491411	-3,2134537	-2,4334002

875 **Table S3. Results of the LMM models for all the data points and for maximum values**
876 **only. Year was set as a fixed factor and Country as a random factor.**
877

Environment	Effect size	p-value
biofilm	4,99E-01	
effluent	1,28E-01	0.0657
feces	1,06E+00	0.224
influent	1,16E-01	0.187
manure	3,90E-02	0.509
sediments	-1,10E-02	0.873
sludge	-3,20E-02	0.710
soil	1,66E-01	0.00566
treated wastewater	-3,96E-01	0.0427
wastewater	-2,60E-01	0.239
water	4,90E-02	0.641

878
879
880

Table S4. ARGs associated with certain environments (only results which are significant after FDR correction). A is the probability that a sample is in the target group (=environment) and B is the probability that you will find the gene in a sample that belongs to this specific environment. Stat is the Indicator value is maximum (i.e. 1) when all occurrences of a gene are found in a single environment and in all samples of this environment.

	Multilevel pattern analysis					
Environment	**ARG**	**A**	**B**	**stat**	**p.value**	**sign**
air	ermX	0.6531	0.8929	0.764	0.000999	***
air	ermC	0.4964	0.7857	0.625	0.000999	***
air	vanB	0.5820	0.5714	0.577	0.000999	***
air	tet32	0.5739	0.5714	0.573	0.000999	***
air	erm36	0.6534	0.5000	0.572	0.000999	***
air	blaGES	0.5228	0.3571	0.432	0.001998	**
air	tetS	0.4605	0.3571	0.406	0.000999	***
air	erm35	0.5465	0.2143	0.342	0.005994	**
manure	aphA1	0.4921	0.3158	0.394	0.00200	**
manure	aadA9	0.4777	0.2895	0.372	0.00400	**
manure	aadA5	0.3539	0.3158	0.334	0.00899	**
sediments	yceLmdtH1	0.6334	0.1429	0.301	0.017	*
sewage	mecA	0.7788	0.7500	0.764	0.000999	***
sewage	blaNDM	0.4203	0.2500	0.324	0.010989	*
air+manure	tnpA1	0.6198	0.5303	0.573	0.000999	***
air+manure	tetL	0.4769	0.2879	0.371	0.004995	**
air+sediments	mexF	0.5955	0.2258	0.367	0.004	**
air+sewage	blaCMY2	0.6082	0.3333	0.45	0.000999	***
air+treated ww	ermA	0.7223	0.3750	0.52	0.000999	***
biofilm+sewage	blaKPC	0.8598	0.4655	0.633	0.000999	***
biofilm+soil	mphC	0.89453	0.09898	0.298	0.017	*

effluent+influent	cfxA	0.8121	0.1517	0.351	0.005	**
feces+sewage	mcr1	0.8190	0.4349	0.597	0.000999	***
influent+sediments	acrA5	0.76078	0.16974	0.359	0.004	**
influent+sediments	mepA	0.59926	0.15498	0.305	0.020	*
manure+sediments	tetR	0.6265	0.2159	0.368	0.003	**
manure+sediments	blaL1	0.9229	0.1057	0.312	0.011	*
manure+sewage	aphA3	0.6397	0.2609	0.409	0.002	**
manure+treated ww	strB	0.5107	0.4800	0.495	0.000999	***
manure+treated ww	strA	0.5850	0.3000	0.419	0.001998	**
manure+wastewater	intI2	0.7724	0.1471	0.337	0.012	*
sediments+sludge	aacC4	0.6827	0.1095	0.273	0.021	*
sediments+treated ww	aadA1	0.6307	0.4129	0.51	0.000999	***
treated ww+wastewater	qnrD	0.5958	0.2381	0.377	0.002	**
air+sediments+soil	oprD	0.8450	0.1543	0.361	0.00699	**
air+sewage+treated ww	ermF	0.6630	0.6458	0.654	0.000999	***
manure+sediments+soil	qacEdelta	0.7794	0.2000	0.395	0.003	**
manure+sediments+wastewater	aadA2	0.6632	0.2451	0.403	0.000999	***
manure+treated ww+wastewater	dfrA1	0.5565	0.2500	0.373	0.002	**
air+influent+sediments+soil	fox5	0.90445	0.12915	0.342	0.00999	**
air+influent+sediments+wastewater	aac6II	0.7777	0.1277	0.315	0.016	*
effluent+influent+manure+sludge	mefA	0.7200	0.2335	0.41	0.003	**
food+influent+sediments+sludge	acrA4	0.91595	0.11512	0.325	0.012	*
influent+manure+sediments+sludge	tnpA2	0.8072	0.1396	0.336	0.00999	**
influent+manure+sediments+sludge	blaOXY1	0.7582	0.1244	0.307	0.01598	*
influent+sediments+sludge+soil	mphA	0.8813	0.1636	0.380	0.00200	**
influent+sediments+sludge+soil	aacC	0.9006	0.1169	0.324	0.00599	**

influent+sediments+sludge+wastewater	tnpA7	0.8550	0.1399	0.346	0.00599	**
influent+sewage+treated ww+wastewater	blaOXA58	0.7708	0.1894	0.382	0.002	**
air+influent+sediments+sludge+soil	oprJ	0.9056	0.1228	0.333	0.00799	**
food+influent+manure+sediments+sludge	tnpA5	0.8363	0.1164	0.312	0.021	*
influent+manure+sediments+sludge+soil	blaSFO	0.9141	0.1115	0.319	0.013	*
air+influnet+sediments+sewage+sludge+wastewater	blaOXA1	0.8080	0.1611	0.361	0.00799	**
air+biofilm+food+influent+manure+sediments+soil	blaampC	0.9005	0.1060	0.309	0.012	*
biofilm+effluent+feces+food+influent+sediments+wastewater	qnrA	0.8600	0.1293	0.333	0.013	*
biofilm+effluent+influent+sediments+sewage+treated ww+wastewater	blaCTXM	0.7535	0.3800	0.535	0.000999	***

891 **Table S5. Comparison of the top most reported and abundant ARGs to existing lists of**
892 **suggested monitoring targets.** Genes in red are already included into clinical monitoring
893 schemes.
894

Antibiotic class	Berendonk et al., 2015	Bengtsson-Palme 2018	Most reported in the current data	Abundant but not often reported
Integrases	intI1		intI1	
Sulfonamide	**sul1**, sul2		**sul1**, sul2	sul3
Beta-lactamases	**blaTEM, blaCTX-M, blaNDM-1, blaVIM, blaKPC**	**blaTEM**	**blaTEM, blaCTX-M, blaSHV**	
	mecA			
Fluoroquinolone	qnrS		qnrS	
Aminoglycoside	**aac-(6')-Ib-cr**	**aac-(6')-Ib**		
	aph	aph(6)-Id		
			strB	
			aadA1	aadA2
Vancomycin	**vanA**			**vanA**
MLSB	**ermB**, ermF	ermC	**ermB**, ermF	
				ereA
Tetracycline	**tetM**	**tetM**, tetQ, tetW, tetC, tetA, tetO,	**tetM**, tetQ, tetW, tetC, tetA, tetO, tetX, tetG, tetB	tetH
Transposase				tnpA4, tnpA5, IS613
Chloramphenicol				floR
Multidrug				mexF
				qacEdelta

895
896
897
898