A global baseline for qPCR-determined antimicrobial resistance gene prevalence across environments

Anna Abramova[1,2,3], Thomas U. Berendonk[4], Johan Bengtsson-Palme[1,2,3]*

[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) in Gothenburg, Sweden

[3] Division of Systems and Synthetic Biology, Department of Life Sciences, SciLifeLab, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden

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

*Corresponding author: johan.bengtsson.palme@chalmers.se

Abstract

The environment is an important component in the emergence and transmission of antimicrobial resistance (AMR). Despite that, little effort has been made to monitor AMR outside of clinical and veterinary settings. Partially, this is caused by a lack of 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 a baseline of AMR in a variety of settings. In an attempt to establish this baseline level, we here performed a comprehensive literature survey, identifying 150 scientific papers containing relevant qPCR data on antimicrobial resistance genes (ARGs) in environments associated with potential routes for AMR dissemination. The collected data

included 1594 samples distributed across 30 different countries and 12 sample types, in a time span from 2001 to 2020. We found that for most ARGs, the typically reported abundances in human impacted environments fell in an interval from 10^{-5} to 10^{-3} copies per 16S rRNA, roughly corresponding to one ARG copy in a thousand bacteria. 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: antibiotic resistance, AMR, qPCR, monitoring, surveillance

55 # Introduction

56 According to the World Health Organization the increasing level of antibiotic resistance over
57 the last decades is a considerable threat to global health[1]. Already, current estimates
58 suggest that hundreds of thousands of deaths are caused by antibiotic resistant bacteria
59 each year worldwide, and this number is expected to reach several millions annually within
60 the next 30 years[2].
61
62 Antibiotic resistance has been mainly investigated in clinical and agricultural settings.
63 However, the role of the environment as a source and dissemination route of resistance has
64 been widely recognized in recent years[3-5]. Antibiotic resistance is common in nature and not
65 limited to human-associated activities, as evidenced by antibiotic resistance genes (ARGs)
66 found in glaciers, permafrost and isolated caves, which have virtually never been visited by
67 humans[6-8]. It is even likely that resistance may be maintained independent of human
68 activities in e.g. soil systems due to the presence of antibiotic-producing microorganisms,
69 including certain fungal taxa[9]. Owing to the extraordinary biological diversity in nature, the
70 external environment can serve as a source of novel ARGs not yet present in clinically
71 relevant pathogens[10]. Furthermore, the environment constitutes a dissemination route for
72 resistant bacteria between humans and animals[11]. However, it is not yet clear to what extent
73 the resistant environmental bacteria and the presence of ARGs in the environment
74 contribute to acquisition and spread of resistance in clinically relevant bacteria. To address
75 these questions, it is essential to identify and quantify resistance in the environment. This in
76 turn would allow pinpointing which environments possess the highest risks to human and
77 domestic animal health and would subsequently enable us to devise measures to prevent –
78 or at least delay – recruitment and dissemination of resistance factors from the environment.
79
80 Monitoring schemes for antibiotic resistance have been developed and put into use in clinical
81 and agricultural settings, but systematic monitoring for antimicrobial resistance (AMR) in the

82 environment is still lacking[12]. Existing monitoring systems largely rely on culturing of
83 pathogens and bacterial indicator species of interest[13], and these approaches may not be
84 directly applicable to the monitoring of resistance in the environment. For example, the
85 majority of environmental bacteria are not culturable using standard laboratory methods and
86 the breakpoints for resistance are adapted for clinically relevant pathogens. Onset of
87 molecular methods, metagenomics and quantitative real-time PCR (qPCR,) allowed to
88 circumvent culturing drawbacks and enabled detection and quantification of individual ARGs
89 in both culturable and unculturable bacteria. Metagenomics allows simultaneous
90 characterization of all ARGs in an environmental sample without *a priori* knowledge, and
91 therefore can be a promising approach for wastewater-based monitoring. However, there are
92 several drawbacks associated with this method, including inherent variability due to lack of
93 standardized protocols, semiquantitative nature of the generated data as well as a relatively
94 poor quantification limit, meaning that rare genes are highly likely to be undetected. In
95 contrast, qPCR can provide very sensitive detection and quantification of ARGs in an
96 environmental sample, which is comparable between different studies and environmental
97 matrices. The major hurdle for using qPCR for monitoring of AMR is the decision of which
98 genes to select as targets in qPCR assays and what reference values to use. However,
99 high-throughput arrays, able to quantify hundreds of ARGs in parallel have been recently
100 developed[14], and can partially alleviate the problem of choosing targets *a priori*[15]. Therefore,
101 qPCR has been recognized as a promising method for AMR monitoring.
102
103
104 To design efficient surveillance strategies, there is a need for comprehensive background
105 data on abundances and prevalence of ARGs occurring in both pristine and human-
106 impacted environments. Establishing background levels of ARGs is a prerequisite for
107 determining “safe levels” for ARGs naturally occurring in the environment and would also
108 provide insights into to which extent an environment has been enriched with ARGs due to
109 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 identify a set of genes and reference abundance values in key environments, which could be used for monitoring purposes. Screening of more than 800 scientific papers for qPCR data for ARGs in various environments, we found that most surveyed environments contained a diverse set of ARGs, but generally at low abundances. 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. Therefore, the available data could not document any significant spatiotemporal changes in AMR, but our analysis of the data identified relative abundances of ARGs that can be suggested as baseline for ARG abundances in the environment.

Materials and Methods

Data collection

Literature collection was performed by mining the PubMed database in April-May 2020, November 2020 and October 2021 (Figure S1). The searches were performed using several strings:

1. ‘"antibiotic resistance genes" AND (monitoring OR surveillance) AND qPCR AND’
2. ‘"antibiotic resistance genes" AND "high throughput" AND’
3. ‘"antibiotic resistance" AND (qPCR OR "quantitative PCR") AND’
4. ‘"antibiotic resistance" AND q-PCR AND‘
5. ‘"antibiotic resistance" AND "quantitative polymerase chain reaction" AND‘
6. ‘"antibiotic resistance" AND "quantitative real-time 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 and animal. Environmental types were selected based on the assessment of potential transmission routes for antibiotic resistance[4,11]. This search resulted in 2636 records, and after removing duplicates in 802 unique papers. Paper suitability was assessed in several steps:

1. Initial assessment by screening the title, abstract, and methods section: review papers, methods papers, as well as papers using only metagenomics as a quantification method were discarded. Only papers using qPCR as a quantification method for environmental samples were retained. In case, papers reported experiments in artificial settings (e.g. bioreactors) as well as with addition of chemical compounds only samples that were not manipulated (e.g. control) were included in the current study.
2. In the second step, we looked through the reported results and supplementary material and selected only studies containing extractable data. Initial search revealed that collected studies used different ways of reporting ARG abundances (e.g. per 16s rRNA, per volume/weight of the sample, per total DNA or per sampled surface). We decided to retain only studies reporting gene abundances per 16S rRNA, or studies reporting numbers which could easily be converted to per 16S rRNA, to enable comparison between different environments (Figure S1). Furthermore, studies reporting only abundances per ARG class but not individual genes were also discarded.
3. For each paper the following information was extracted: first author, DOI, year, country, sample description (e.g. distance from a contamination source, season, type of fertilizer applied, stage of wastewater treatment process, source etc.) and abundances of antibiotic resistance genes (ARGs). For several studies missing information about the exact year of data collection, we extrapolated the sampling

162 year based on the average time required from the data collection to the publication
163 inferred from the studies reporting specific sampling dates, corresponding to an
164 estimate of two years.

165 4. We divided the collected samples into several “sample types”: “air”, “biofilm”, “feces”
166 (containing both animal and human feces), “manure”, “food”, “sediments”, “soil”,
167 “water”, including water from different sources such as potable, surface or reclaimed
168 water etc.; “wastewater”, consisting of wastewater, sewage and influent samples,
169 because these samples have a similar/overlapping origin; “sludge” which includes
170 semi-solid by-products of different stages of the wastewater treatment process; and
171 “swab” corresponding to samples collected by swabbing the washing machines,
172 shower drains and dishwashers surfaces with a sterile cotton swab.

173 5. We further classified collected samples according to exposure to anthropogenic
174 impact: “impacted” when a sample was obviously affected by human activity such as
175 effluent samples or wastewater, “likely impacted” when it was not directly affected by
176 an obvious pollution source but was located in a populated area, “likely unimpacted”
177 if the sample was collected from relatively “pristine” environments, “feces/manure”
178 and “unknown”. The “unknown” category included samples which were clearly
179 impacted by human activity, but where the activity aims at removing bacteria, e.g.
180 drinking/tap water.

181 6. The ARG abundances were collected from text, tables and figures. Information from
182 figures was extracted using WebPlotDigitizer v4.3
183 (https://automeris.io/WebPlotDigitizer). Abundances were reported as log10 ARG
184 abundance per 16S rRNA (Supporting Information S1).

185 Data analysis

186 Data were analyzed in RStudio v1.3.1073[16] with the R package ggplot2 v3.3.5[17] used to plot
187 the graphs. To determine the abundance distributions for ARGs across environments,

188 density plots of the ARG abundances were used, and the R summary function was used to
189 find the quartiles for each ARG. To investigate if there were any changes in ARG levels over
190 time, we built a linear mixed model (LMM). LMMs allow incorporation of not only fixed effects
191 but also random effects to account for cluster-correlated data from distinct sources of
192 variability[18]. The model was designed with ARG *Abundance* as a response variable and
193 sampling *Year* as explanatory variable and included sampling *Country* as a random effect:
194
195 $$Abundance = Year + (1|Country) + \varepsilon$$
196
197 Analyses were performed using lme4 v1.1-27.1[19] and lmerTest v3.1-3[20] packages.
198
199 To identify a set of suitable qPCR targets, we investigated how well the data from a subset
200 of genes could predict the abundance of the other genes in the data matrix by performing a
201 correlation analysis. Only genes with more than 10 entries were included in this analysis.
202 The R package stats (version 3.6.2) *cor* function with "pairwise.complete.obs" was used to
203 calculate Spearman correlations.

204 Results and Discussion

205 Current data on AMR in the environment is biased

206 Mining of the PubMed database resulted in 802 unique papers and revealed a bias in the
207 existing literature towards a few already fairly well studied environments (Figure S2).
208 Unsurprisingly, the majority of the identified papers matched environmental categories that
209 have been already extensively studied as primary sources of antibiotics and/or antibiotic
210 resistance to the environment. These include environments linked to wastewater treatment
211 plants, hospitals and industrial facilities, as well as agriculture and livestock production. This
212 also highlights the undersampled nature of the other environments. Among them are

213 categories associated with human mobility. Human mobility, in particular international travel,
214 could significantly contribute to dissemination of ARGs across the globe[10,21-23]. Despite that,
215 categories associated with travel activity, such as “public transport” and “airports”, were
216 represented by just a handful of papers. Another poorly investigated category is water
217 associated with recreational activities. It has been suggested that accidental ingestion of
218 water during, for example, swimming can serve as a potential route for resistant
219 environmental bacteria into the human gut[24].
220
221 In total, we identified 150 studies containing 1594 samples providing relevant information on
222 levels of ARGs measured by qPCR. These samples represented 12 different sample types
223 with “water”, “feces” and “sediments” being the most common (Figure 1B). The samples
224 originated from 30 different countries (Figure 1A), with China being the most studied among
225 them, both in terms of number of studies and sample types. The distribution of studies
226 highlighted white spots on the map; in particular in South America and just a handful of
227 studies from Africa. This is worrisome, as South America and Africa are among the regions
228 with the highest rates of self-medication and supply of non-prescribed antibiotics in the
229 world[25,26]. In 2014, a WHO report on clinical data outlined Africa’s lack of established AMR
230 surveillance systems[1] and a review from 2017 reported that more than a third of the
231 countries on the continent did not have recent AMR data published in the public domain, and
232 that only little of the reported data were from surveillance[27]. The scarcity of public data on
233 AMR in general is corroborated by the current literature survey. The majority of samples in
234 the selected studies were collected between 2015-2018, with a gradual increase in the
235 number of publications over the last years (Figure 1C).
236

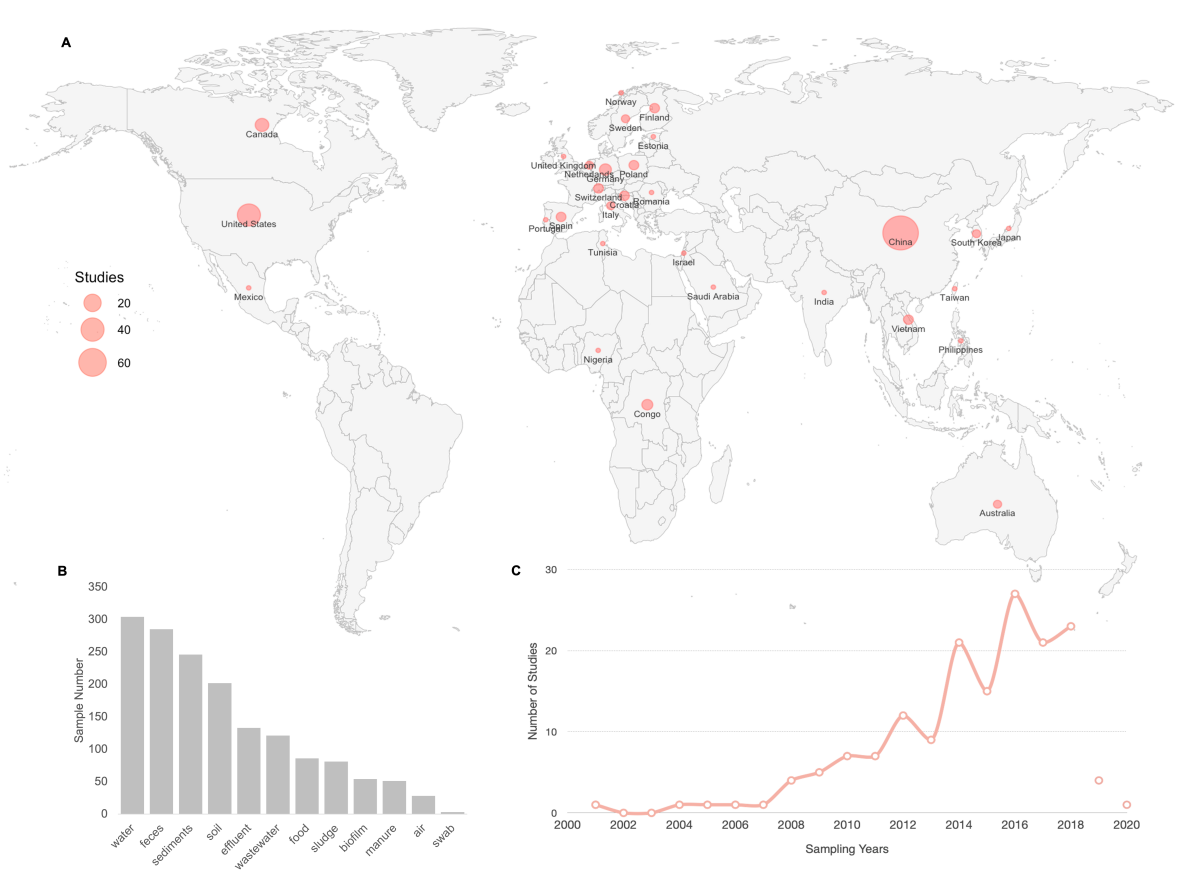
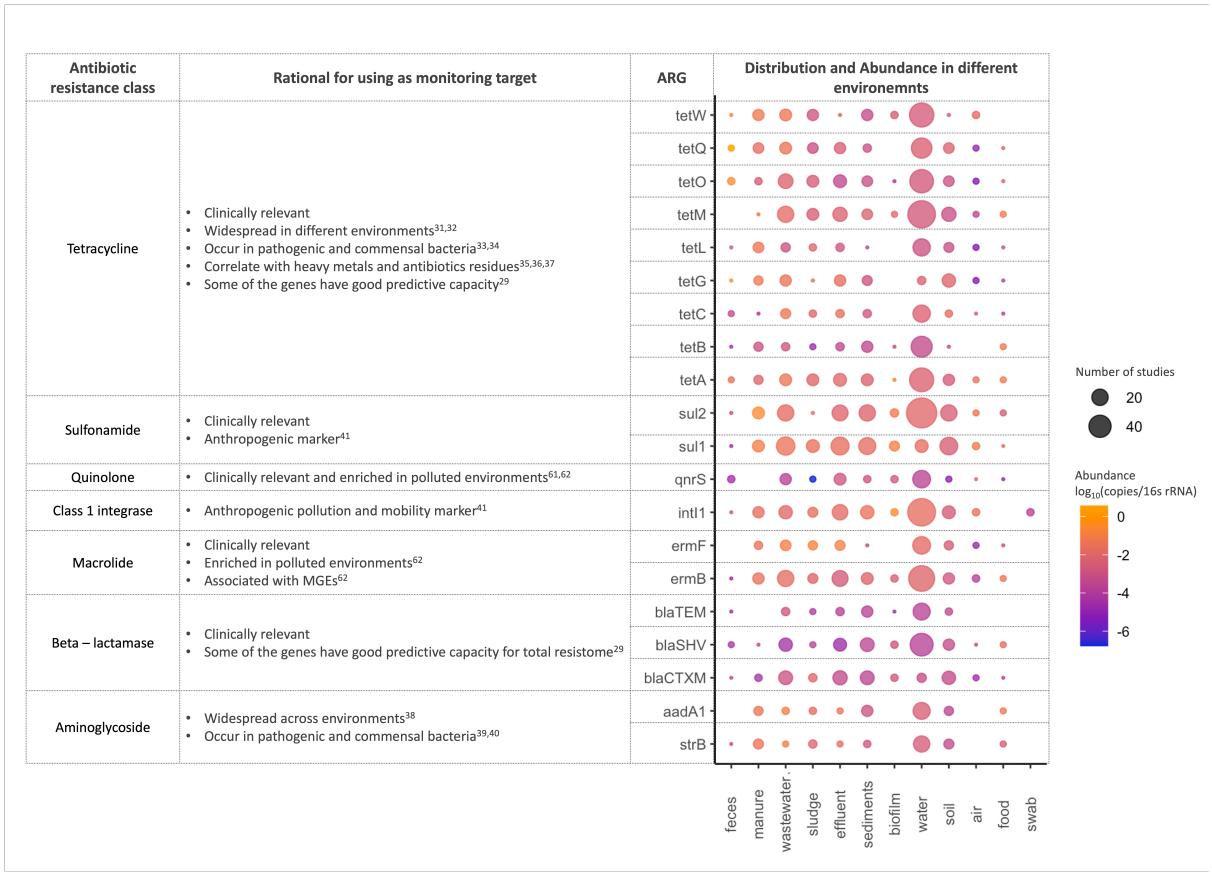


Figure 1. Overview of the collected data: A) geographical distribution of surveyed studies; B) number of samples representing different sample types; C) time distribution of collected studies (years 2019 and 2020 are represented by dots not connected by the line due to studies containing data from these years are yet to be published).

Monitoring targets for AMR surveillance in the environment

It is generally agreed that the choice of targets for environmental AMR surveillance should be governed by the purpose of monitoring, which can include *i)* assessment of risks for transmission of already resistant pathogens to humans via environmental routes, *ii)* analysis of general levels of resistance in the human population, and *iii)* monitoring the evolution of resistance in the environment to identify emerging resistance threats[28]. There are several

251 previous studies suggesting a number of ARGs targets for environmental monitoring which
252 typically include clinically relevant ARGs to capture genes of immediate concern to human
253 health, anthropogenic pollution markers, as well as mobility markers. The list of most often
254 reported targets in the current survey was not very different from the previously suggested
255 targets in Berendonk et al., 2015[4], Bengtsson-Palme 2018[29] and Keenum et al., 2021[30]
256 (Figure 2).
257



258

259 **Figure 2 Most reported qPCR targets and their proposed value as monitoring targets**
260 **suggested in existing literature; MGEs stands for Mobile Genetic Elements.**
261

262 Among the ARG targets commonly used in environmental qPCR studies, many genes have
263 already been included in clinical AMR testing (Figure 2) (e.g.
264 https://www.oxfordbiosystems.com/AMR, https://www.opgen.com/acuitas/amr-gene-panel/)
265 or provide resistance to antibiotics of high clinical concern listed by WHO[41]. These include

266 clinically relevant β-lactamases such as bla_{CTX-M}, bla_{TEM} and bla_{SHV}, the sulfonamide
267 resistance gene *sul1,* as well as *ermB* and *tetM* conferring resistance to macrolides and
268 tetracyclines, respectively. Although these genes are already common both in the
269 environment and among human pathogens, they could be useful to predict general
270 resistance levels or transmission risks. However, their usefulness in environmental
271 monitoring for emerging threats is somewhat questionable. At the same time, they may
272 function as a gauge of the total antibiotic resistance content in an environment, as some of
273 them have been shown to be good predictors of the total resistome, particularly *tetM* and
274 bla_{TEM}[29].
275
276 Not typically included in routine clinical monitoring but previously suggested as a marker for
277 anthropogenic AMR dissemination[42] in the environment is the class 1 integrase *intI1* gene.
278 Despite that this gene is not a resistance gene itself, it is often linked to ARGs. According to
279 the existing literature, it is widespread in the environment, found in both pathogenic and
280 commensal bacteria, enriched by anthropogenic activities and correlates well with the
281 abundances of other commonly occurring ARGs[43]. Interestingly, despite being so
282 widespread, our data showed that this gene has lower abundances in the likely unimpacted
283 sampling locations, such as arctic soils and alpine lakes, supporting its usefulness as an
284 indicator of anthropogenic pollution in the environment. Similarly, the sulfonamide resistance
285 gene *sul2* has also been suggested as a monitoring target by Berendonk et al. (2015)[4], was
286 among the most reported in our study and was widespread across all the environments.
287 However, *sul2* added little information compared to *intI1* (Figure S3), so these two genes
288 would be mostly redundant in AMR monitoring.
289 The plasmid-mediated fluoroquinolone resistance gene *qnrS* was among the most reported
290 genes in our study. It was found in all environments except manure and swabs. Contrary to
291 what has been reported previously[61,62], we did not detect an enrichment of this gene in
292 polluted environment specifically. As it is the only commonly reported fluoroquinolone ARG,
293 it is an interesting target for monitoring. However, it has the drawback of generally not being

294 able to induce clinically relevant levels of resistance without additional resistance
295 mechanisms, which makes its clinical impact somewhat limited.
296 Among the most common target genes, several tetracycline resistance genes showed
297 significantly higher abundances in fecal/manure samples than in other human-impacted
298 environments (including *tetA*, *tetG*, t*etH*, *tetO*, *tetQ* and *tetW*). Furthermore, these genes are
299 among the most powerful ARGs for predicting the diversity and abundance of other ARGs[29]
300 and can therefore, despite their widespread distribution, still be useful in monitoring to
301 extrapolate other parts of the resistome.
302
303 There were also several genes that were not often included in qPCR studies, but when they
304 were, they often appeared in relatively high abundances (above one copy per 1000 bacteria)
305 (Figure S5). Such ARGs could be interesting additional monitoring targets in future AMR
306 monitoring schemes. Among these genes was the vancomycin resistance gene *vanA,* a
307 previously suggested indicator of antibiotic resistance contamination of clinical origin, which
308 is thought to be uncommon in the environment[44]. In our data, however, *vanA* was about as
309 abundant as other ARGs in the environment when it has been looked for. Another gene that
310 seemed to correlate with anthropogenic activities is *ereA*, a macrolide resistance gene which
311 has previously been reported to be the most abundant in metal polluted soil[45] and is
312 enriched by long-term application of manure[46]. In contrast, the *mexF* gene has been
313 reported from many different environments, including soil, sediments and water[21] and is
314 suggested to naturally occur in unaffected environments such as pristine Antarctic soils[47],
315 indicating that it might be a useful target for identifying enrichment of ARGs occurring in
316 environmental microbial communities irrespectively of pollution from fecal material. The *sul3*
317 gene has been rarely reported, despite being rather abundant in the cases when it was
318 quantified. According to previous studies, *sul3* is typically less frequent and abundant than
319 *sul1* and *sul2*[31,48]. Similarly to *sul1* and *sul2*, it has been shown to be associated with class 1
320 integrons and has been detected on a conjugative plasmid, suggesting its potential to be
321 horizontally transferred and disseminated in the environment[49]. Interestingly, *sul3* was first

322 detected in an *E. coli* isolate from a pig, and it was later found in both healthy and diseased
323 humans. Furthermore, it is often enriched in polluted environments and is present in both
324 commensal and pathogenic bacteria[50]. Several other genes (*tnpA4, tnpA5, qacEdelta,*
325 *aadA2, floR, tet32, cmlA1* and *mefA*), which were not often reported, but often highly
326 abundant when detected, would also be potentially suitable additional targets for AMR
327 monitoring (Figure S5).
328

329 Around one in a thousand environmental bacteria carry clinical ARGs

330 In the current study, we estimated typical abundance levels of ARGs in different
331 environments. We found that for most ARGs, the typically reported abundances fell in an
332 interval from 10^{-5} to 10^{-3} copies per 16S rRNA, roughly corresponding to one ARG copy in a
333 thousand bacteria (Figure 3B). Importantly, there was a strong bias towards studies of
334 environments already impacted by humans (e.g. WWTPs, agriculture and animal
335 production), and therefore this range rather reflects ARG levels in environments already
336 affected by human activity.
337
338 That said, most of the ARGs measured in likely unimpacted sampling locations had lower
339 typical abundances, below 10^{-5} copies per 16S rRNA (Figure 3A). In particular, samples of
340 water and soil from the arctic tundra showed the lowest ARG abundances. There are
341 previous studies suggesting that ARG abundances fluctuate with latitude, which to some
342 degree could be explained by optimal growth temperature for the microbes carrying these
343 ARGs[51]. Furthermore, there is evidence for a distance-decay relationship for abundance of
344 ARGs on the global and continental scales[52,53], which could explain higher abundances of
345 ARGs in the water from alpine lakes in comparison to water from arctic tundra. Overall,
346 however, it is hard to draw any major conclusions due to the small number of samples from
347 relatively unimpacted locations.
348

349 Our data revealed distinct distributions of several tetracycline genes (including *tetA*, *tetG*,
350 t*etH*, *tetO*, *tetQ* and *tetW*) in feces/manure samples (Figure 3A, Figure S6). Tetracycline
351 ARGs constituted the predominant class among the most reported qPCR targets and are
352 typically associated with human fecal samples[54]. Notably, several of these genes showed
353 distinct abundance distributions for impacted environments and feces/manure samples,
354 suggesting that they can be used as markers for human/animal fecal contamination in the
355 environment. Particularly, *tetA* and *tetG* show distinct distributions in impacted and non-
356 impacted environments and could serve as indicators for anthropogenic pollution, while *tetW*
357 shows overlapping distributions making it less useful to ascertain human impact. Similarly,
358 the clinically relevant $bla_{CTX\text{-}M}$ and bla_{TEM} genes also showed a separation in impacted and
359 non-impacted environments, although that was less clear than for tetracycline genes.
360
361 The “unknown” category was characterized by relatively high abundances of ARGs and for
362 some of the genes even higher levels than in impacted samples. This category comprised
363 samples of water and biofilm from households (unspecified), as well as potable and
364 reclaimed water distribution systems. We classified these samples as “unknown” since they
365 are obviously impacted by human activity, but the impact targets the removal of ARGs and
366 bacteria. Despite that, ARGs in some of these samples had higher relative abundances than
367 in the fecal/manure samples (Figure 3A, Figure S6). These samples may yet not have higher
368 abundances of ARGs per amount of sample, as this would also depend on how much
369 bacteria are present in the samples per volume or weight.

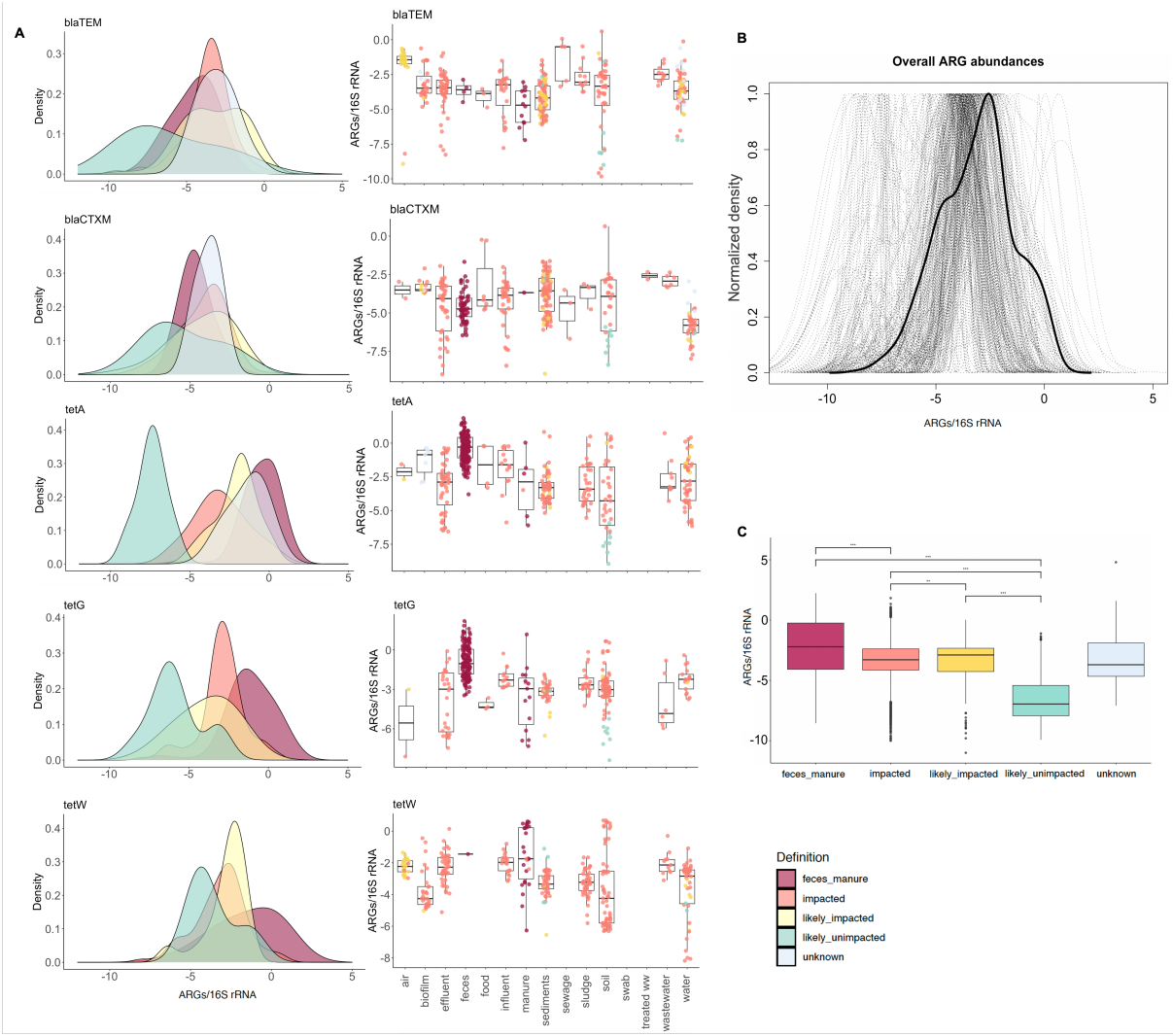


Figure 3. Relative abundance distributions of surveyed ARGs. A) Density plots represent abundance ranges of blaTEM, blaCTXM, tetA, tetG and tetW, in different environmental types divided according to exposure to anthropogenic impact (see material and methods for the detailed description), and corresponding bar plots show what samples contributed to each distribution. B) Density plot representing relative abundance distributions across all the studied genes with the bold line indicating the average distribution. C) Difference between abundance ranges in different environmental types. The environmental types are color-coded as follows: “feces/manure” in purple, “impacted” in orange, “likely impacted” in yellow, “likely unimpacted” in green and “unknown” type in blue. Note that abundances are given relative to the number of 16S rRNA copies, which means that these

distributions do not necessarily correspond to the total exposure to resistant bacteria in a given environment.

It is unclear whether ARG abundances have increased over time

To investigate if there were changes in ARG abundances over time, we used linear mixed models on gene abundances from each of the environments, accounting for variability between countries (Figure S4, Table S2). Due to the successively increased sensitivity of the methods used for qPCR 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 per year reported from a particular sample type 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 produce the estimates), except for in effluent, sediments, sludge, wastewater and water where the trends were negative (Table S2). That said, the trends were significant only for biofilm, effluent, food, sediments, sludge and wastewater. 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.

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 2 and 3). To explore if they would be good predictors of overall ARG abundances, we performed a correlation analysis of the most reported genes and the rest of the genes in the data matrix (Figure S3). This analysis revealed a certain degree of redundancy in terms of what information can be gained from the proposed monitoring targets. ARGs proposed as monitoring targets that strongly correlate with each other convey similar information on ARG

407 abundance and diversity and may therefore not be very useful to use in the same panel for
408 ARG monitoring in the environment. Genes that were overall redundant included the *tetO*
409 and *tetA* genes, *strB* and *ermF*, $bla_{CTX\text{-}M}$ and bla_{SHV}, *tetG* and *tetM*, as well as *intI1* and *sul2*
410 (Figure S3). If resources are constrained, it would seem wise not to use several genes in
411 these smaller groups together in the same ARG panel for monitoring.
412

413 Specific uses of qPCR as a monitoring tool

414 In this study, we have exclusively focused on the use of qPCR as a tool for monitoring the
415 abundance of ARGs in the environment. In general, the qPCR technique has several
416 benefits, but also some drawbacks as a surveillance tool. An important advantage is that it is
417 a highly sensitive method that can detect much lower levels of ARGs than is currently
418 possible using shotgun metagenomic sequencing[55]. Furthermore, it can operate on very
419 minute quantities of DNA, which makes it suitable also for low-biomass samples. In addition,
420 as it is performed on DNA, it can detect ARGs in non-culturable bacteria and is functional
421 also on complex samples with many different species.
422 However, application of qPCR as an AMR monitoring tool requires *a priori* knowledge on a
423 set of targets as well as how to interpret the results in terms of when an ARG occurrence
424 pattern becomes a concern. The limiting factor of predefined targets can be partially
425 overcome by using qPCR arrays with hundreds of genes[15], but this also comes with
426 increased costs and may not (at present) be feasible for large-scale routine monitoring of
427 environmental AMR. Despite that, even a handful of selected targets can still provide useful
428 information about the total resistome situation, fecal contamination and potentially HGT
429 intensity. From a monitoring perspective, the background ARG levels identified in this study
430 could be used to infer an increase in ARG abundances as a sign of pollution and/or
431 selection. Based on the collected data, we advise the use of 10 times the third quartile (3Q)
432 values for a given ARG (Table S1) to determine the upper limit for what should be viewed as
433 a deviation from normal background levels in the environment. A more fundamental aspect

434 of such deviations is whether an increase in the abundance of ARGs or mobile genetic
435 elements in a particular environment or at a specific time point is a relevant indicator of a
436 selective pressure for resistance. Importantly, an increase of ARG abundances without any
437 context is more likely to be an indicator of human pollution[56], but it cannot be ruled out that
438 such a change could be due to a specific selection pressure from antibiotics[57], by co-
439 selection from other antibacterial compounds[58], or simply from taxonomic shifts that are
440 unrelated to antibiotic resistance.
441 In the end, the suitability of qPCR for environmental AMR surveillance comes down to the
442 purpose of monitoring and what type of actions one might want to take based on the
443 monitoring results. If detecting dissemination of known high-risk ARGs is the sole purpose of
444 monitoring, qPCR is an excellent method thanks to its sensitivity. However, if identification of
445 emergent resistance threats is the goal, qPCR is unlikely to give useful guidance; instead,
446 shotgun metagenomics[59] or selective culturing followed with genetic profiling would provide
447 more useful information.
448

449 Outcomes and recommendations

450 In this study, we performed a literature survey to explore the abundance and prevalence of
451 ARGs in various environments as quantified by qPCR. We found that, overall, previous
452 suggestions for ARGs to be included in environmental AMR monitoring[4,29] seem relevant.
453 Particularly, inclusion of the *intI1*, *sul1*, bla_{TEM}, $bla_{CTX\text{-}M}$ and *qnrS* genes in environmental
454 monitoring seems essential, along with a selection of the tetracycline genes, in particular at
455 least one of *tetA* or *tetG,* which could serve as fecal pollution markers. However, there are
456 also genes that are not often looked for in qPCR surveys that perhaps should be, including
457 *sul3*, *vanA*, *tetH*, *aadA2*, *floR*, *ereA* and *mexF*. These genes are abundant in some
458 environments, but were not often included in qPCR studies of environmental AMR. We also
459 provide environmental baseline levels for the ARGs studied through qPCR (Figure 3, Table
460 S1); for most ARGs the typical relative abundance falls in an interval from 10^{-5} to 10^{-3}. It

should be noted that this is the range of normal abundances of ARGs and should not be considered a maximum acceptable limit of these ARGs in any given environment. Such maximum acceptable levels need to be determined taking risks to human health (and potentially the environment) as well as the numbers of bacteria in a given volume of sample into account[5,60], and would also need to consider transmission routes to humans[11]. The different standards of reporting DNA abundances constituted a complicating factor for this study. Particularly, the choice to report the abundances of ARGs relative to the 16S rRNA gene, which was necessitated by the availability of data, limits the interpretability of our data in terms of exposure risks. The absence of clear trends of increases or decreases in ARG abundances over time indicates a need for more systematic time series data on ARG abundances in a variety of environments. Our results also highlight the scarcity of AMR data from parts of the world, particularly from Africa and South America, and underscores the need for a concerted effort to quantify typical background levels of AMR in the environment more broadly to enable efficient environmental AMR surveillance schemes akin to those that exist in clinical settings.

Acknowledgments

JBP acknowledges funding from the Swedish Research Council (VR; grant 2019-00299) under the frame of JPI AMR (EMBARK; JPIAMR2019-109), the Data-Driven Life Science (DDLS) program supported by the Knut and Alice Wallenberg Foundation (KAW 2020.0239), the Swedish Foundation for Strategic Research (FFL21-0174), 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 Marcus Wenne, Emil Burman, Sebastian Wettersten, Tora Hulterström and anonymous reviewers for constructive feedback on earlier versions of 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., Manaia, C.M., Merlin, C., Fatta-Kassinos, D., Cytryn, E., Walsh, F., Bürgmann, H., Sørum, H., Norström, M., and Pons, M.-N., *Tackling antibiotic resistance: the environmental framework.* Nature Reviews Microbiology, 2015. **13**(5): p. 310-317.
5. Larsson, D.J., Andremont, A., Bengtsson-Palme, J., Brandt, K.K., de Roda Husman, A.M., Fagerstedt, P., Fick, J., Flach, C.-F., Gaze, W.H., and Kuroda, M., *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 Petrova, M., *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., Wang, W., Koteva, K., Barton, H.A., McArthur, A.G., and Wright, G.D., *A diverse intrinsic antibiotic resistome from a cave bacterium.* Nature communications, 2016. **7**(1): p. 1-10.
8. Van Goethem, M.W., Pierneef, R., Bezuidt, O.K., Van De Peer, Y., Cowan, D.A., and Makhalanyane, T.P., *A reservoir of 'historical'antibiotic resistance genes in remote pristine Antarctic soils.* Microbiome, 2018. **6**(1): p. 1-12.
9. Bahram, M., Hildebrand, F., Forslund, S.K., Anderson, J.L., Soudzilovskaia, N.A., Bodegom, P.M., Bengtsson-Palme, J., Anslan, S., Coelho, L.P., and Harend, H., *Structure and function of the global topsoil microbiome.* Nature, 2018. **560**(7717): p. 233-237.
10. Bengtsson-Palme, J. and Larsson, D.J., *Antibiotic resistance genes in the environment: prioritizing risks.* Nature Reviews Microbiology, 2015. **13**(6): p. 396-396.
11. Bengtsson-Palme, J., Kristiansson, E., and Larsson, D.J., *Environmental factors influencing the development and spread of antibiotic resistance.* FEMS microbiology reviews, 2018. **42**(1): p. fux053.
12. Pruden, A., Vikesland, P.J., Davis, B.C., and de Roda Husman, A.M., *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., Pruden, A., Alvarez, P.J., Aga, D., Bürgmann, H., Li, X.-d., Manaia, C.M., Nambi, I., Wigginton, K., and Zhang, T., *Toward a comprehensive strategy to mitigate dissemination of environmental sources of antibiotic resistance*. 2017, ACS Publications.
14. Stedtfeld, R.D., Guo, X., Stedtfeld, T.M., Sheng, H., Williams, M.R., Hauschild, K., Gunturu, S., Tift, L., Wang, F., Howe, A., Chai, B., Yin, D., Cole, J.R., Tiedje, J.M.,

and Hashsham, S.A., *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., Jameel, S., Ali, J., Saleem Ur Rehman, H., Tauseef, I., Farooq, U., Jamal, A., and Ali, M.I., *Contributions and Challenges of High Throughput qPCR for Determining Antimicrobial Resistance in the Environment: A Critical Review.* Molecules, 2019. **24**(1): p. 163.
16. Team, R., *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., Donaldson, L., Correa-Cano, M.E., Evans, J., Fisher, D.N., Goodwin, C.E., Robinson, B.S., Hodgson, D.J., and Inger, R., *A brief introduction to mixed effects modelling and multi-model inference in ecology.* PeerJ, 2018. **6**: p. e4794.
19. Bates, D., Mächler, M., Bolker, B., and Walker, S., *Fitting linear mixed-effects models using lme4.* arXiv preprint arXiv:1406.5823, 2014.
20. Kuznetsova, A., Brockhoff, P.B., and Christensen, R.H., *lmerTest package: tests in linear mixed effects models.* Journal of statistical software, 2017. **82**(1): p. 1-26.
21. D'Souza, A.W., Boolchandani, M., Patel, S., Galazzo, G., van Hattem, J.M., Arcilla, M.S., Melles, D.C., de Jong, M.D., Schultsz, C., and Dantas, G., *Destination shapes antibiotic resistance gene acquisitions, abundance increases, and diversity changes in Dutch travelers.* Genome medicine, 2021. **13**(1): p. 1-21.
22. Frost, I., Van Boeckel, T.P., Pires, J., Craig, J., and Laxminarayan, R., *Global geographic trends in antimicrobial resistance: the role of international travel.* Journal of travel medicine, 2019. **26**(8): p. taz036.
23. Von Wintersdorff, C.J., Penders, J., Van Niekerk, J.M., Mills, N.D., Majumder, S., Van Alphen, L.B., Savelkoul, P.H., and Wolffs, P.F., *Dissemination of antimicrobial resistance in microbial ecosystems through horizontal gene transfer.* Frontiers in microbiology, 2016. **7**: p. 173.
24. Leonard, A.F., Zhang, L., Balfour, A.J., Garside, R., Hawkey, P.M., Murray, A.K., Ukoumunne, O.C., and Gaze, W.H., *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.
25. Auta, A., Hadi, M.A., Oga, E., Adewuyi, E.O., Abdu-Aguye, S.N., Adeloye, D., Strickland-Hodge, B., and Morgan, D.J., *Global access to antibiotics without prescription in community pharmacies: a systematic review and meta-analysis.* Journal of Infection, 2019. **78**(1): p. 8-18.
26. Yeika, E.V., Ingelbeen, B., Kemah, B.L., Wirsiy, F.S., Fomengia, J.N., and van der Sande, M.A., *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.
27. Tadesse, B.T., Ashley, E.A., Ongarello, S., Havumaki, J., Wijegoonewardena, M., González, I.J., and Dittrich, S., *Antimicrobial resistance in Africa: a systematic review.* BMC infectious diseases, 2017. **17**(1): p. 1-17.
28. Huijbers, P.M.C., Flach, C.F., and Larsson, D.G.J., *A conceptual framework for the environmental surveillance of antibiotics and antibiotic resistance.* Environ Int, 2019. **130**: p. 104880.
29. Bengtsson-Palme, J., *The diversity of uncharacterized antibiotic resistance genes can be predicted from known gene variants—but not always.* Microbiome, 2018. **6**(1): p. 125.
30. Keenum, I., Liguori, K., Calarco, J., Davis, B.C., Milligan, E., Harwood, V.J., and Pruden, A., *A framework for standardized qPCR-targets and protocols for quantifying antibiotic resistance in surface water, recycled water and wastewater.* Critical Reviews in Environmental Science and Technology, 2022. **52**(24): p. 4395-4419.

31. Suzuki, S., Ogo, M., Takada, H., Seki, K., Mizukawa, K., Kadoya, A., Yokokawa, T., Sugimoto, Y., Sato-Takabe, Y., and Boonla, C., *Contamination of antibiotics and sul and tet (M) genes in veterinary wastewater, river, and coastal sea in Thailand.* Science of The Total Environment, 2021. **791**: p. 148423.
32. Wu, N., Qiao, M., Zhang, B., Cheng, W.-D., and Zhu, Y.-G., *Abundance and diversity of tetracycline resistance genes in soils adjacent to representative swine feedlots in China.* Environmental Science & Technology, 2010. **44**(18): p. 6933-6939.
33. Bryan, A., Shapir, N., and Sadowsky, M.J., *Frequency and distribution of tetracycline resistance genes in genetically diverse, nonselected, and nonclinical Escherichia coli strains isolated from diverse human and animal sources.* Applied and environmental microbiology, 2004. **70**(4): p. 2503-2507.
34. Fairchild, A., Smith, J., Idris, U., Lu, J., Sanchez, S., Purvis, L., Hofacre, C., and Lee, M., *Effects of orally administered tetracycline on the intestinal community structure of chickens and on tet determinant carriage by commensal bacteria and Campylobacter jejuni.* Applied and environmental microbiology, 2005. **71**(10): p. 5865-5872.
35. Lu, L., Liu, J., Li, Z., Zou, X., Guo, J., Liu, Z., Yang, J., and Zhou, Y., *Antibiotic resistance gene abundances associated with heavy metals and antibiotics in the sediments of Changshou Lake in the three Gorges Reservoir area, China.* Ecological Indicators, 2020. **113**: p. 106275.
36. Chen, B., Liang, X., Huang, X., Zhang, T., and Li, X., *Differentiating anthropogenic impacts on ARGs in the Pearl River Estuary by using suitable gene indicators.* Water research, 2013. **47**(8): p. 2811-2820.
37. Huang, Y.-H., Liu, Y., Du, P.-P., Zeng, L.-J., Mo, C.-H., Li, Y.-W., Lü, H., and Cai, Q.-Y., *Occurrence and distribution of antibiotics and antibiotic resistant genes in water and sediments of urban rivers with black-odor water in Guangzhou, South China.* Science of the total environment, 2019. **670**: p. 170-180.
38. Kheiri, R. and Akhtari, L., *Antimicrobial resistance and integron gene cassette arrays in commensal Escherichia coli from human and animal sources in IRI.* Gut Pathogens, 2016. **8**(1): p. 40.
39. Kiiru, J., Butaye, P., Goddeeris, B.M., and Kariuki, S., *Analysis for prevalence and physical linkages amongst integrons, ISEcp1, ISCR1, Tn21 and Tn7 encountered in Escherichia coli strains from hospitalized and non-hospitalized patients in Kenya during a 19-year period (1992–2011).* BMC Microbiology, 2013. **13**(1): p. 109.
40. Fonseca, É.L., dos Santos Freitas, F., Scheidegger, É.M.D., Jacinto, T., and Vicente, A.C.P., *Class 2 integrons in multidrug-resistant Acinetobacter baumannii circulating in different Brazilian geographic regions.* International journal of antimicrobial agents, 2011. **38**(1): p. 95-96.
41. Gillings, M.R., Gaze, W.H., Pruden, A., Smalla, K., Tiedje, J.M., and Zhu, Y.-G., *Using the class 1 integron-integrase gene as a proxy for anthropogenic pollution.* The ISME journal, 2015. **9**(6): p. 1269-1279.
42. World Health Organization, *The 2019 WHO AWaRe classification of antibiotics for evaluation and monitoring of use*. In The 2019 WHO AWaRe classification of antibiotics for evaluation and monitoring of use, 2019.
43. Zhang, A.N., Li, L.-G., Ma, L., Gillings, M.R., Tiedje, J.M., and Zhang, T., *Conserved phylogenetic distribution and limited antibiotic resistance of class 1 integrons revealed by assessing the bacterial genome and plasmid collection.* Microbiome, 2018. **6**(1): p. 130.
44. Narciso-da-Rocha, C., Varela, A.R., Schwartz, T., Nunes, O.C., and Manaia, C.M., *blaTEM and vanA as indicator genes of antibiotic resistance contamination in a hospital–urban wastewater treatment plant system.* Journal of global antimicrobial resistance, 2014. **2**(4): p. 309-315.
45. Chen, J., Li, J., Zhang, H., Shi, W., and Liu, Y., *Bacterial heavy-metal and antibiotic resistance genes in a copper tailing dam area in northern China.* Frontiers in microbiology, 2019. **10**: p. 1916.

46. Li, J., Xin, Z., Zhang, Y., Chen, J., Yan, J., Li, H., and Hu, H., *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.
47. Wang, F., Stedtfeld, R.D., Kim, O.-S., Chai, B., Yang, L., Stedtfeld, T.M., Hong, S.G., Kim, D., Lim, H.S., and Hashsham, S.A., *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.
48. Hoa, P.T.P., Nonaka, L., Viet, P.H., and Suzuki, S., *Detection of the sul1, sul2, and sul3 genes in sulfonamide-resistant bacteria from wastewater and shrimp ponds of north Vietnam.* Science of the Total Environment, 2008. **405**(1-3): p. 377-384.
49. Antunes, P., Machado, J., and Peixe, L., *Dissemination of sul3-containing elements linked to class 1 integrons with an unusual 3' conserved sequence region among Salmonella isolates.* Antimicrobial agents and chemotherapy, 2007. **51**(4): p. 1545-1548.
50. Hammerum, A.M., Sandvang, D., Andersen, S.R., Seyfarth, A.M., Porsbo, L.J., Frimodt-Møller, N., and Heuer, O.E., *Detection of sul1, sul2 and sul3 in sulphonamide resistant Escherichia coli isolates obtained from healthy humans, pork and pigs in Denmark.* International journal of food microbiology, 2006. **106**(2): p. 235-237.
51. Tourna, M., Freitag, T.E., Nicol, G.W., and Prosser, J.I., *Growth, activity and temperature responses of ammonia-oxidizing archaea and bacteria in soil microcosms.* Environmental Microbiology, 2008. **10**(5): p. 1357-1364.
52. Zheng, D., Yin, G., Liu, M., Chen, C., Jiang, Y., Hou, L., and Zheng, Y., *A systematic review of antibiotics and antibiotic resistance genes in estuarine and coastal environments.* Science of the Total Environment, 2021. **777**: p. 146009.
53. Bahram, M., Hildebrand, F., Forslund, S.K., Anderson, J.L., Soudzilovskaia, N.A., Bodegom, P.M., Bengtsson-Palme, J., Anslan, S., Coelho, L.P., and Harend, H., *Structure and function of the global topsoil microbiome.* Nature, 2018. **560**(7717): p. 233-237.
54. Zhang, A.-N., Gaston, J.M., Dai, C.L., Zhao, S., Poyet, M., Groussin, M., Yin, X., Li, L.-G., van Loosdrecht, M., and Topp, E., *An omics-based framework for assessing the health risk of antimicrobial resistance genes.* Nature communications, 2021. **12**(1): p. 1-11.
55. Lundström, S.V., Östman, M., Bengtsson-Palme, J., Rutgersson, C., Thoudal, M., Sircar, T., Blanck, H., Eriksson, K.M., Tysklind, M., and Flach, C.-F., *Minimal selective concentrations of tetracycline in complex aquatic bacterial biofilms.* Science of the Total Environment, 2016. **553**: p. 587-595.
56. Karkman, A., Pärnänen, K., and Larsson, D.J., *Fecal pollution can explain antibiotic resistance gene abundances in anthropogenically impacted environments.* Nature communications, 2019. **10**(1): p. 1-8.
57. Bengtsson-Palme, J., Boulund, F., Fick, J., Kristiansson, E., and Larsson, D.G.J., *Shotgun metagenomics reveals a wide array of antibiotic resistance genes and mobile elements in a polluted lake in India.* Frontiers in Microbiology, 2014. **5**.
58. Pal, C., Bengtsson-Palme, J., Kristiansson, E., and Larsson, D.G.J., *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.
59. Yang, D., Heederik, D.J., Scherpenisse, P., Van Gompel, L., Luiken, R.E., Wadepohl, K., Skarżyńska, M., Van Heijnsbergen, E., Wouters, I.M., Greve, G.D. and Jongerius-Gortemaker, B.G., *Antimicrobial resistance genes aph (3')-III, erm (B), sul2 and tet (W) abundance in animal faeces, meat, production environments and human faeces in Europe.* Journal of Antimicrobial Chemotherapy, 2022. *77*(7): p.1883-1893.
60. Pruden, A., Larsson, D.G.J., Amézquita, A., Collignon, P., Brandt, K.K., Graham, D.W., Lazorchak, J.M., Suzuki, S., Silley, P., Snape, J.R., Topp, E., Zhang, T., and Zhu, Y.-G., *Management Options for Reducing the Release of Antibiotics and*

Antibiotic Resistance Genes to the Environment. Environmental Health Perspectives, 2013. **121**(8): p. 878-885.

61. Castrignanò, E., et al., *Enantiomeric profiling of quinolones and quinolones resistance gene qnrS in European wastewaters.* Water research, 2020. **175**: p. 115653.

62. 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.