bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The global groundwater resistome: core ARGs and their dynamics - an *in silico* re-analysis of publicly available groundwater metagenomes

Ioannis D. Kampouris[1,2,*], Thomas U. Berendonk[1], Johan Bengtsson-Palme[3,4,5], Uli Klümper[1,*]

[1] Institute of Hydrobiology, Technische Universität Dresden, 01217 Dresden, Germany

[2] Arctic Research Centre, Department of Biology, Aarhus University, Aarhus, Denmark

[3] Division of Systems Biology, Department of Biology and Biological Engineering, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden

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

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

* Corresponding authors

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

16 **Corresponding authors:**

17 Ioannis D. Kampouris (ORCID: 0000-0003-2093-5930) & Uli Klümper (ORCID: 0000-0002-

18 4169-6548)

19 Technische Universität Dresden

20 Institute of Hydrobiology

21 Zellescher Weg 40

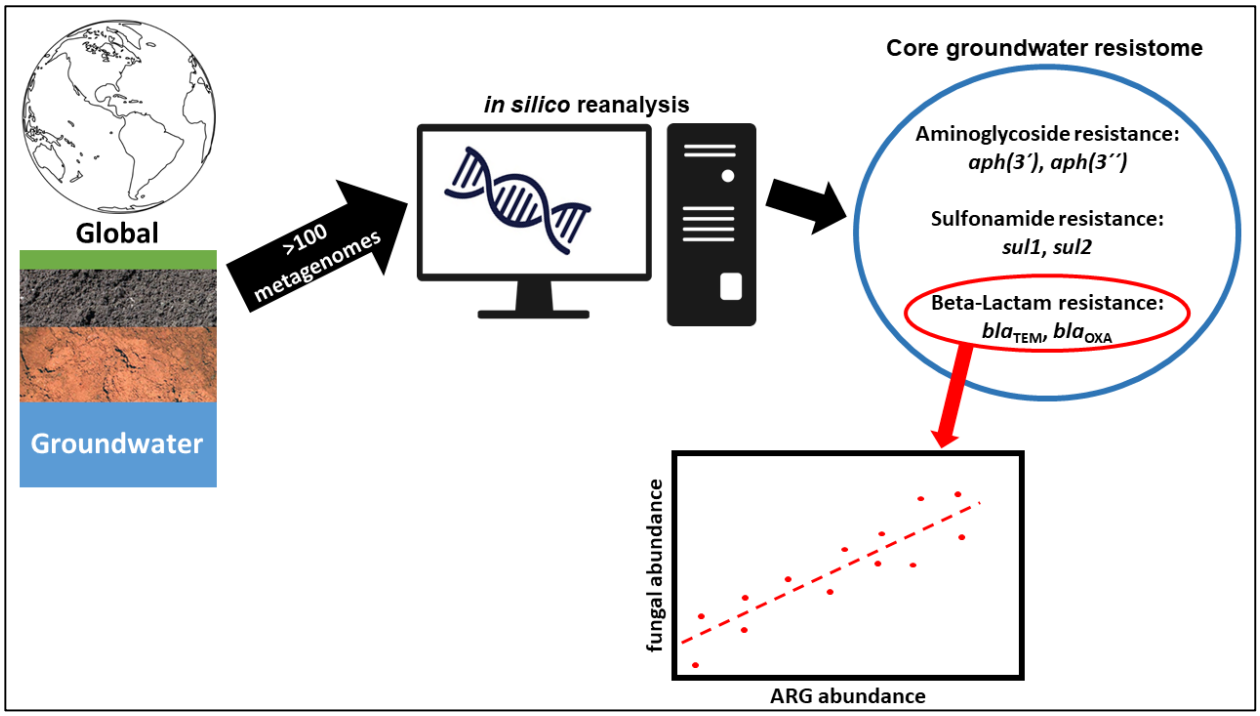
22 01062 Dresden

23 Germany

24 E-mail: ioannisds@yahoo.gr; uli.kluemper@tu-dresden.de

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Graphical abstract



bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Despite the importance of groundwater environments as drinking water resources, there is currently no comprehensive picture of the global levels of antibiotic resistance genes in groundwater. Moreover, the biotic and abiotic factors that might shape the groundwater resistome remain to be explored on a global scale. Herein, we attempted to fill this knowledge gap by *in silico* re-analysis of publicly available global groundwater metagenomes. We first investigated the occurrence of antibiotic resistance genes (ARGs) to define the core groundwater resistome. We further tested whether the ARG dissemination in the pristine groundwater environments could be explained by natural ecological processes such as competition between fungal and bacterial taxa. Six ARGs encoding resistance to aminoglycosides (*aph(3')*, *aph(3'')*), sulfonamides (*sul1*, *sul2*), and β-lactams (bla_{OXA}, bla_{TEM}) occurred in at least 50% of samples at high abundance, thereby constituting the core groundwater resistome. ARG abundances differed significantly between countries and only weakly correlated with bacterial community composition. While only limited effects of anthropogenic impacts could be observed, ecological interactions played a significant role in shaping the abundance patterns of at least a number of the core ARGs. Fungal abundance positively correlated with bla_{TEM} and bla_{OXA} abundance, ARGs that confer resistance to β-lactams, regularly produced by fungi. However, no direct correlation was determined for the remainder of the core ARGs. Still, using co-occurrence network analysis we identified that the fungal abundance acted as a hub-node that included bla_{OXA} and bla_{TEM}, but also indirectly contributed to the abundance of aminoglycoside ARG *aph(3')*. Hence, interactions between bacteria and fungi including potential antibiotic production can contribute to the dissemination of ARGs in groundwater environments. Consequently, fungal/bacterial SSU ratio could serve as an indicator for the abundance of certain ARGs in the pristine groundwater environments.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Keywords

Groundwater; Antimicrobial resistance; Antibiotic resistance; Fungi; Metagenomic re-analysis

Highlights

- Core GW resistome included *aph(3')*, *aph(3'')*, *sul1*, *sul2*, *bla*$_{OXA}$ and *bla*$_{TEM}$
- Limited effects of anthropogenic impacts on GW resistome
- Fungal/bacterial abundance positively correlated with *bla*$_{TEM}$ and *bla*$_{OXA}$ abundance
- Fungal/bacterial abundance can serve as indicator for certain ARGs in groundwater

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1. Introduction

The global rise in antimicrobial resistance (AMR) represents a major threat to future human health (Laxminarayan el., 2013). Tackling it requires a “One Health” approach that considers AMR dynamics and proliferation between the human, veterinary and environmental spheres (Hernando-Amado et al., 2019). Drinking water resources provide one of the immediate connections between environmental and human microbiomes (Vaz-Moreira et al., 2014). Among these, groundwater (GW) ecosystems constitute the most common freshwater and drinking water resource in the majority of the world (Szekeres et al., 2018; Griebler and Avramov et al., 2015; Herrmann et al., 2019). GW environments are characterized by high microbial diversity and complexity (Griebler and Lueders, 2009; Flynn, et al., 2013; Griebler and Avramov et al., 2014), with GW microbiota playing important roles in several biogeochemical cycles (Flynn et al., 2013; Sonthiphand et al., 2019). Due to the role of GW environments as a major drinking water resource, understanding the occurrence of AMR in GW environments is highly relevant for tackling AMR through a “One Health” approach (Hernando-Amado et al., 2019).

While several studies have focused on the importance of GW microbiota in biogeochemical cycles (Flynn et al., 2013; Sonthiphand et al., 2019; Retter et al., 2021) or in their response to pollution with toxic compounds (Taş et al., 2018; Sonthiphand et al., 2019), only few have looked into the occurrence dynamics of ARGs in GW using qPCR or metagenomic approaches (Szekeres et al., 2018; Zhang et al., 2019; Zaouri et al., 2020). The potential anthropogenic impact on AMR in GW was demonstrated for GW beneath a commercially operated wastewater irrigated field (Kampouris et al., 2022). Here the abundance of specific antibiotic resistance genes (ARGs) increased in accordance with the infiltration of the respective antibiotics from wastewater into the GW. However, the majority of ARG abundance dynamics in more pristine GW environments remains difficult to explain. Consequently, a global and comprehensive

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

picture of the natural ARG levels in GW and the non-anthropogenic factors that might shape the GW resistome is needed.

Such global studies have been performed in terrestrial, non-anthropogenically impacted environments, such as soil and surface marine waters, and generally linked ARG dissemination partly to the competition between fungal and bacterial taxa (Bahram et al., 2018). Fungi regularly thrive in soils, in close interaction with other biota (Bahram et al., 2018) and can manipulate and shape the indigenous bacterial communities (Johnston et al., 2019). For example, several fungal taxa produce β-lactam antibiotics (e.g. penicillin) (Aly et al., 2011). Consequently, the specific complex fungi-bacteria interactions have been theorized as the cause underlying the natural prevalence of β-lactam ARGs in the environment (e.g. occurrence of bla_{TEM} and bla_{CTX-M} variants in relatively pristine soils) (Gatica et al., 2015). In GW environments most of the detected fungi function as saprophytic organisms, enabling the degradation of organic matter and performing organic carbon recycling (Nawaz et al., 2018). However, how the presence of these fungi and the resulting fungal-bacteria interactions in the humid, dark and pristine GW environments could affect the GW resistome has not yet been fully explored.

Herein, we aimed to fill the knowledge gaps regarding which ARGs constitute the core global GW resistome and if, similar to in pristine soils, interactions with fungi could provide an explanatory variable in shaping it. To this end, we performed an *in silico* re-analysis of publicly available global GW metagenomes retrieved from the NCBI sequencing read archive (SRA), specifically investigating which genes constituted the core resistome and how they related to the overall taxonomy of the GW communities.

2. Methodology

2.1 Data collection of groundwater metagenomes

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Public metagenome datasets for samples from global GW environments were searched and obtained from the NCBI sequencing read archive (SRA). The search queries included the terms "groundwater", "aquifer", or "subsurface water", for the matrices; and "shotgun sequencing" or "wgs" for the sequencing method. The information from SRA was linked to publications and locations, whenever available. Accession numbers and linked publications for all the retrieved metagenomic datasets (100 metagenomes in total) are given in Table S1. An additional 30 identified candidate metagenomes from peer-reviewed studies were unfortunately not made publicly available or did not pass the quality criteria presented in the next section and were thereby excluded from the study.

2.2 Annotation of antibiotic resistance gene profile and taxonomical composition

For each metagenomic dataset general quality control and trimming were performed with the tool cutadapt (v3.1, Martin, 2011), with the following command: cutadapt --cores=10 --cut 20 -q 10 --minimum-length 90 --max-n 0 --max-ee 0.1. The selection of a maximum expected error (ee) of 0.1 allowed only high quality sequences to pass. Sequences with a length of less than 90 bp were filtered out, to ensure a sufficient read length for ARG annotation. ResFinder (Version 4), a database of mobile, acquired antibiotic resistance genes (Bortolaia et al., 2020) was translated from nucleotide sequences into amino acid sequences using Biopython (Cock et al., 2009). ARGs were annotated against the translated ResFinder database using the command "blastx" in DIAMOND (Buchfink et al., 2015) with the following parameters: minimum identity 99%, minimum match length 30 amino acids. The parameters were chosen to be conservative to reduce false positive hits. In case of paired-end sequencing, matches on the second paired read were counted only if there was no match on the first read. The tool METAXA2 (version 2.2.3) (Bengtsson-Palme et al., 2015) was used for the identification of total small subunits of ribosome (SSU), 16S rRNA for prokaryotes and 18S rRNA for eukaryotes, to determine taxonomic composition, using the default settings. Screening for crAssphage sequences, an

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

indicator for anthropogenic fecal pollution, (Karkman et al., 2019) was performed with "ngless" (Coelho et al., 2019), which utilizes a version of the BWA-MEM algorithm for alignment (Li, 2013; Li and Durbin, 2010).

To exclude any potential effects of differing sequencing depth on the estimated abundance of ARGs, we performed a correlation of total ARG abundances with total bacterial counts. This proved non-significant (Spearman rank correlation, R=-0.17, *p*=0.1, Fig. S1B), hence sequencing depth can be excluded as a confounding factor.

2.3 Data analyses and statistics

Following ARG annotation and determination of the taxonomic composition, the results were analyzed in R (v4.0, R Core Team, 2019). The total bacterial and fungal counts for each metagenomic sample were calculated with the "tidyverse" packages (v1.0.4, Wickham, 2019). The ARG, bacterial and fungal relative abundances were calculated similarly using the same packages. The fungal 18S to bacterial 16S rRNA ratio was calculated using the "mutate" function from the package "dplyr" (v.1.0.10, Wickham et al., 2022). The package "ggplot2" (v.3.3, Wickham, 2016) was used for graphical representations.

Differences in the ARG composition based on Euclidean distance were visualized and evaluated using the "vegan" package (v.2.5.6, Oksanen et al., 2019) by generation of NMDS plots and statistical PERMANOVA tests. ARGs that were present with less than two reads in less than four metagenomes from a single country were removed from the differential analysis for ARGs. Countries with less than three available, high-quality metagenomes were excluded from the location based analysis as well. All data was $\log_{10}$-transformed. Since the sequencing depth of the retrieved metagenomes differed, for estimation of the differential gene abundances and distance metrics we first calculated the limit of detection (LOD) of the different samples (Fig. S1A). Then, zeros in abundance were replaced with an abundance of 10^{-8} gene/SSU, which

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

was one order of magnitude below the sample with lowest LOD ($3x10^{-7}$). The differences in bacterial community composition were calculated similarly, with the sole exception that it was based on the Bray-Curtis dissimilarity of bacterial taxa at the family level.

For comparing the differential abundance of every single ARG per location, the Kruskal-Wallis test was performed with the use of the package "ggpubr" (v. 0.2.2, Kassambara, 2019). Mantel and Procrustes tests between ARG profile (Euclidean distance) and bacterial community composition (Family level, Bray-Curtis distance) were performed with the "vegan" package.

For correlation analyses, the data for different bacterial taxonomical groups, ARGs and fungal/bacterial 16S rRNA ratio was $\log_{10}$ transformed and Spearman correlations coefficients were estimated with the package "ggpubr". Samples with less than two positive hits for specific taxonomical groups or ARGs were excluded from the correlation analysis. In addition, linear mixed-effect models (package "lme4", v1.1.3 Bates et al., 2022) were performed to account for confounding variability in sampling, DNA extraction, etc., to subsequently verify the hypothesized correlations. In these linear mixed-effect models, we used the original study of each metagenome as random variable.

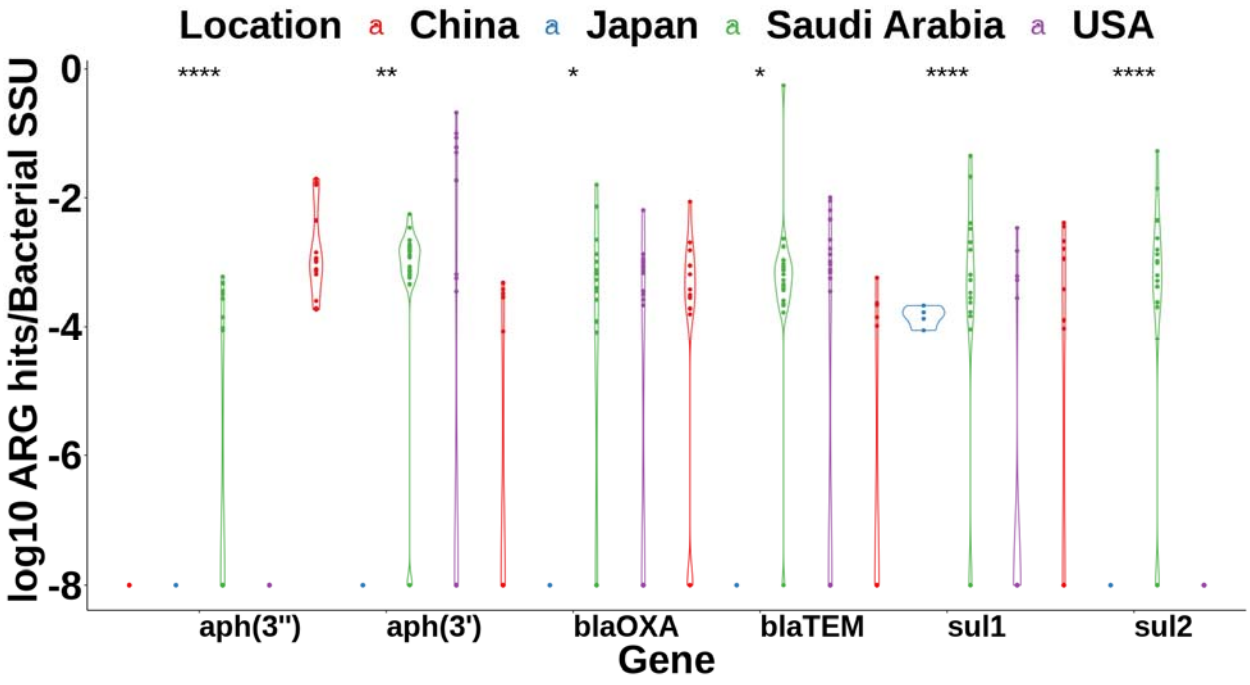
To reveal a) whether total fungal abundance correlates with changes in bacterial community composition and b) whether these correlations can be linked to the fungal/ARG correlations, a co-occurrence network was constructed using Spearman correlation and Benjamini-Hochberg correction, with a threshold of $p<0.05$. Samples without any positive hit were excluded to avoid correlations due to zero inflated data. For inclusion in the co-occurrence network, a minimum threshold was set: 25 samples with positive hits for each ARG or phylogenetic group. The co-occurrence network was constructed with the packages "igraph" (Csardi et al., 2005) and "ggraph" (v2, Pedersen, 2022).

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

3. Results and Discussion

3.1 The core groundwater resistome

In total 99 of the 100 screened metagenomes (Table S1) from diverse geographical locations, including the US, Saudi Arabia, Japan and Germany, exceeded the high quality criteria (read size <90 bp, expected error rate <0.1/read) for subsequent re-analysis. ARGs were successfully detected in 87 of the 99 metagenomes. Overall, the common global GW resistome consisted of 24 ARGs which were detected in at least three metagenomes at abundances above 10^{-5} hits per bacterial SSU (Fig. S2). These confer resistance to 13 antibiotic classes including aminoglycosides, β-lactams, sulfonamides, and macrolides. Among these 24 ARGs, only the sulfonamide ARGs *sul1* and *sul2*, the β-lactam ARGs bla_{OXA} and bla_{TEM} and the aminoglycoside ARGs *aph*(3') and *aph*(3'') occurred in at least 50% of the metagenomes and throughout displayed significantly higher relative abundances compared to the remaining ARGs (Kruskal-Wallis test $p<2.2\times10^{-16}$, Fig. S2). They hence constitute the core GW resistome (Fig. 1).



bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 1. Relative abundance of the ARGs that comprise the core resistome (occurred in at least 50% of the metagenomes). Countries with less than three high-quality groundwater metagenomes in which ARGs were detected, were excluded from this analysis. Significant differences were assessed with Kruskal-Wallis test (*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001, n=4-30). Values of $\log_{10}$≤-8 ARG hits/Bacterial SSU represent samples with ARGs below the limit of detection, rather than actual values.

Among the observed core ARGs, *bla*$_{TEM}$ abundance was consistently higher, compared to other β-lactam ARGs. Variants of *bla*$_{TEM}$ have regularly been found to occur in high abundance in soil microbiota, with no clear relation to anthropogenic influence (Gatica et al., 2013; Kampouris et al., 2021; Wang et al., 2022). In previous studies, levels of *bla*$_{TEM}$ were found to be similar between wastewater and GW environments (Kampouris et al., 2021), while *bla*$_{TEM}$ was the dominant β-lactam ARG in GW environments, other β-lactam ARGs displayed up to two orders of magnitude higher abundances than *bla*$_{TEM}$ in wastewater (Kampouris et al., 2022). Similar trends were observed when comparing pristine and agricultural soils in Germany (Kampouris et al., 2021) and in China (Wang et al., 2022).

3.2 Antibiotic resistance gene profiles diverge between different countries

Resistome profiles based on the 24 detected common ARGs grouped significantly based on the originating countries (Fig. 2A, PERMANOVA test, Euclidean Distance, R^2=0.33, p=1x10^{-6}, n=4-30; sample number differed per-study). Abundances of most ARGs strongly depended on location: for example, the highest abundance for most ARGs was detected in GW metagenomes originating from Saudi Arabia (Fig. 1). This was especially true for those ARGs that commonly occur in high abundance in wastewater microbiomes, such as *sul1* and *sul2* (Caucci et al., 2016; Cacace et al., 2019) (Fig. 1, Kruskal Wallis, p<0.001, n=4-30). These two genes confer resistance to sulfonamides, antibiotics of synthetic origin that have previously been shown to accumulate in GW with parallel increase of sulfonamide ARGs, especially in locations with extensive wastewater reuse for irrigation purposes (Avisar et al., 2009; Kampouris et al., 2022). Indeed, rates of wastewater reuse in Middle Eastern countries such as Saudi

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Arabia far exceed those in the other countries tested here (Jones et al., 2021; Liao et al., 2021). Consequently, the direct infiltration of antibiotic resistant bacteria from wastewater irrigation, or the infiltration of selective agents such as sulfonamide antibiotics could explain the increased rates of ARGs in GWs of Saudi-Arabia. However, this hypothesis needs to be further tested, since the herein analyzed metagenomes might have originated from sampling different depths and types of GW environment (e.g. geyser or enclosed aquifer; Table S1), which could have acted as a confounding variable on the differences in ARG profiles across the varying locations.

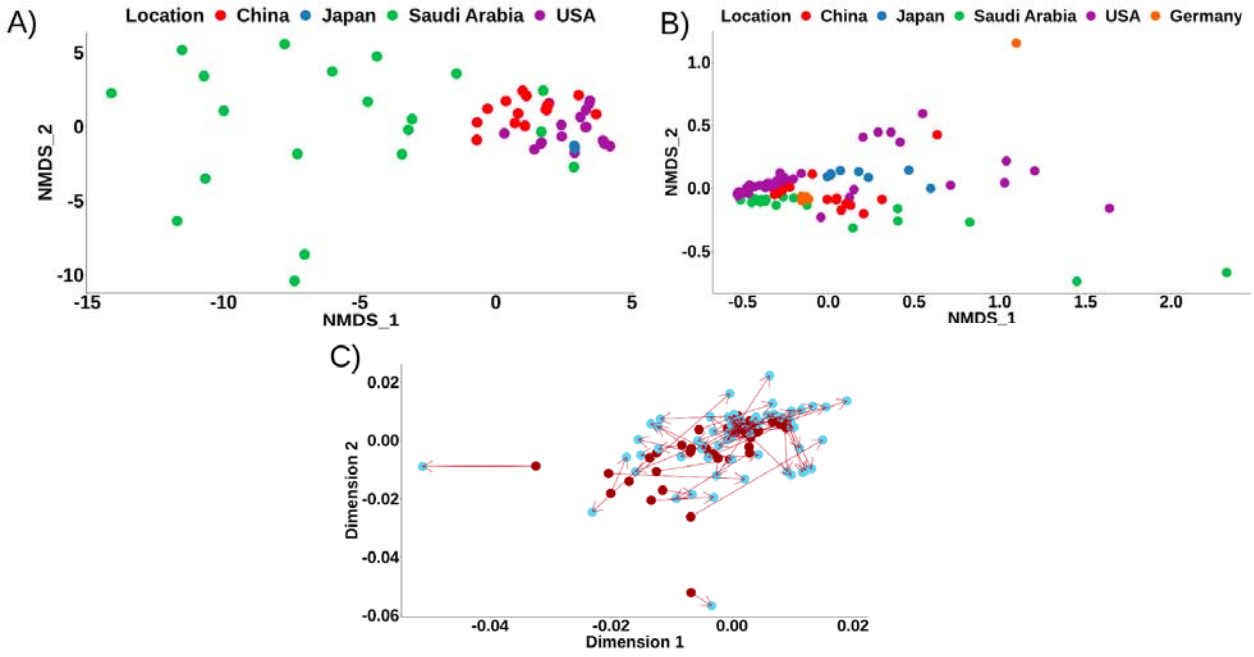


Figure 2. A) NMDS grouping of ARG profiles of groundwater metagenomes by Euclidian distance based on country of origin (PERMANOVA test, Euclidean Distance, R^2=0.24, *p*=1x10^{-6}, n=4-30). ARGs that did not occur in more than three metagenomes from at least one single country were excluded. Countries with less than three high-quality groundwater metagenomes in which ARGs were detected, were also excluded from this analysis. B) Bacterial community composition based clustering of groundwater metagenomes by Bray-Curtis dissimilarity based on country of origin (PERMANOVA test, Bray-Curtis Distance, R^2=0.33, *p*=1x10^{-6}, n=6-41). C) Procrustes rotation plot between ARG profiles (Euclidean distance) and bacterial community composition (Bray-Curtis distance). Procrustes rotation was used to rotate the dissimilarity matrix of the bacterial community composition (β-diversity, Weighted Unifrac Distance) to maximum similarity with the target

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

dissimilarity matrix of the ARG composition (Euclidean distance) by minimizing the sum of squared differences. The Procrustes plot visualizes the association between bacterial community and ARG composition. The length of the arrows visualizes the degree of match between the two ordinations following Procrustes rotation (arrow-start: β-diversity, arrow-end: ARG composition). Mantel test, Spearman correlation *rho*=0.25, *p*=0.001, Procrustes test *rho*=0.62, n=69.

To determine if such a potential direct effect of infiltration of fecal microorganisms to these GW environments exists, we quantified the abundance of crAssphage in the samples, which has been suggested as an indicator of pollution with fecal anthropogenic microorganisms (Karkman et al., 2019). Consequently, crAssphage presence would indicate that wastewater derived organisms were the main driving force underlying increased ARG abundance in these GW environments. However, no crAssphage reads were detected in any of the studied metagenomes, indicating that the infiltration of fecal organisms can be excluded as an explanatory variable for the increased levels of *sul1* and *sul2*. Still, the infiltration of selective agents independent of fecal organisms remains an option that has previously been observed for certain GW environments (Kampouris et al., 2022). However, this could not be tested in this study due to the lack of associated metadata on concentrations of antibiotics.

Similar to ARG profiles, the bacterial community compositions clustered by countries (Fig. 2B, PERMANOVA test, Bray-Curtis distance, R^2=0.24, p=10^{-6}, n=6-41; sample number differed per-study). Still, bacterial community composition dissimilarity provided only a minor explanation for ARG compositional dissimilarities as only a weak significant correlation was found (Mantel test, Spearman correlation *rho*=0.25, *p*=0.001, Procrustes test, *rho* = 0.62, n=69, Fig. 2C).

3.3 Correlation of fungal and antibiotic resistance gene abundances in groundwater metagenomes

Aside from bacterial community composition, we aimed at further exploring the underlying drivers of resistome diversity and abundance in the GW microbiota by evaluating if ecological

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

interactions with natural producers of antibiotics such as fungi and *Actinobacteria* could play a role in ARG dissemination (Bahram et al., 2018). Fungal activity has indeed been hypothesized to contribute to ARG dissemination and maintenance in environments with low levels of anthropogenic pollution with bacteria or selective agents (Bahram et al., 2018). We hence evaluated the correlation of the six core GW ARGs with fungal relative abundance (fungal/bacterial SSUs in the metagenomes). A clear correlation between fungal per bacterial abundance and bla_{TEM} abundance (Spearman rho=0.48, p=0.0039, Fig. 3) and a weak but significant correlation for bla_{OXA} abundance (Spearman rho=0.36, p=0.049, Fig. 3) were observed. No correlation was detected for the remaining core GW ARGs (*sul1, sul2, aph*(3'), *aph*(3''), p>0.05). Consequently, fungal abundance correlated mainly with the levels of bla_{TEM} and bla_{OXA}, which confer resistance to β-lactam antibiotics, commonly produced by several fungal species as secondary metabolites (Nesme and Simonet, 2015). The observed correlations for bla_{TEM} and bla_{OXA} were further verified using a linear mixed model. Here, the original study that the metagenomes were derived from was set as a random effect variable to counter potential study based biases (ARG Rel. Abundance ~ Fungal Rel. Abundance + 1|Original_Study) (p=0.0152). When reducing study based biases with the linear mixed model the previously barely significant correlation for bla_{OXA} (p=0.049) clearly increased in significance (p=0.008).

We further examined the correlation of ARG abundance with *Actinobacteria,* known as the major group of bacterial antibiotic producers (Miao et al., 2010). *Actinobacteria* and bla_{TEM} abundance weakly correlated initially (p=0.047, Fig. S3), but this could not be confirmed using the linear mixed model (p>0.05, ARG Rel. Abundance ~ *Actinobacteria* Rel. Abundance + 1|Original_Study). None of the remaining core-resistome ARGs significantly correlated with *Actinobacteria* abundance (p>0.05, Fig. S3).

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

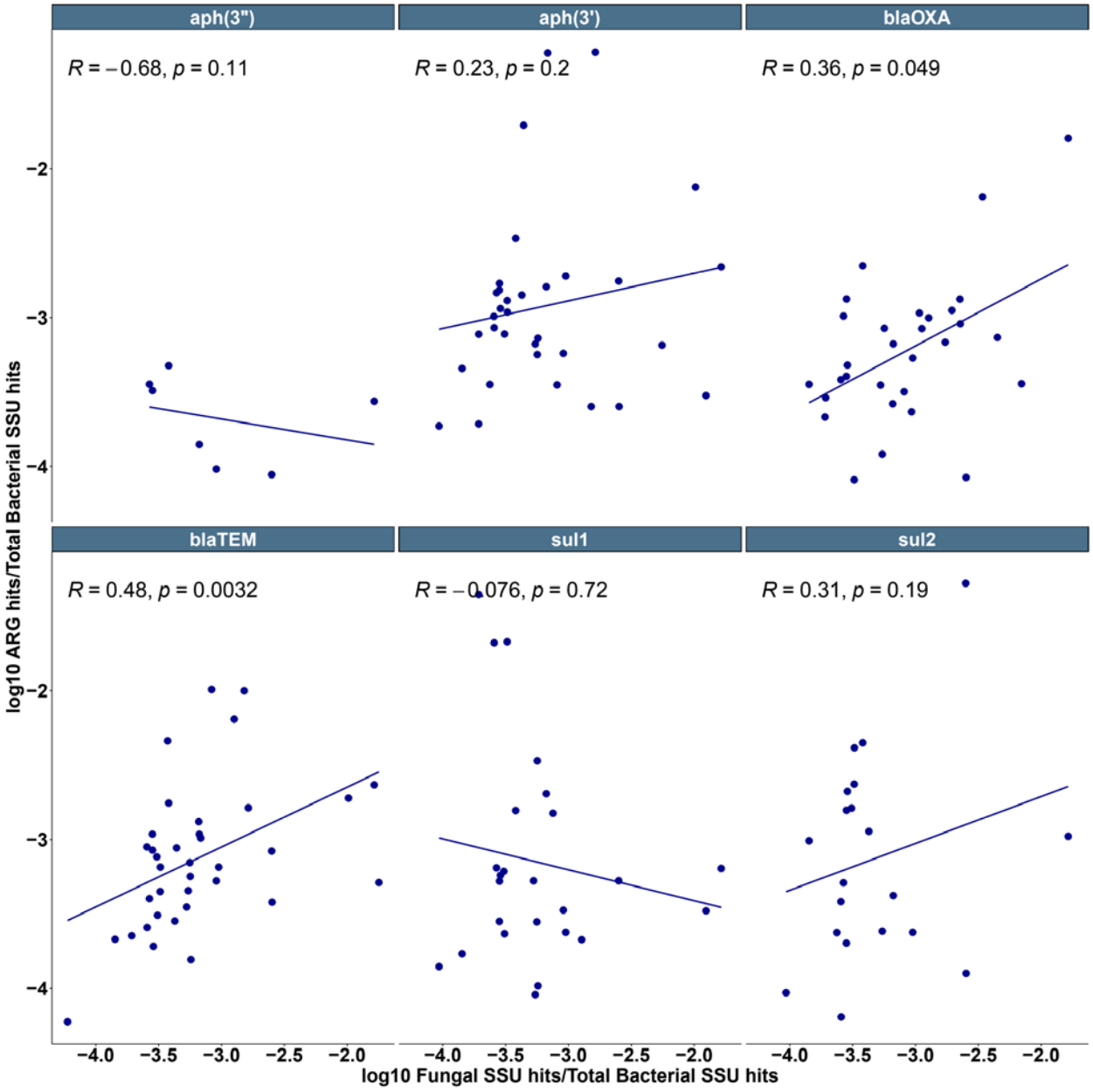


Figure 3. Linear regression and correlation of the relative abundance of individual core groundwater ARGs to ribosomal fungal/bacterial small subunit (SSU) ratio (Spearman rank-correlation). Samples without fungal or ARG counts were excluded.

3.4 Fungal abundance might serve as an indicator for ARG abundance in groundwater environments

To further explore if fungal relative abundance can explain ARG GW dynamics, co-occurrence network analysis (Spearman correlation, p<0.05, Benjamini-Hochberg correction) with bacterial

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

297 community composition (lowest taxonomical level: Order), fungi/bacteria SSU ratio and ARG
298 abundances was performed. Of all 24 ARGs tested only three of the core GW ARGs *bla*$_{TEM}$,
299 *bla*$_{OXA}$ and *aph*(3') were included as a part of the correlation network (Fig. S4). As two of these
300 were already previously associated with positive correlations with fungal abundance we
301 extracted all correlations from the network which either one of the ARGs or the fungal/bacterial
302 SSU ratio were a part of (Fig. 4). The three ARGs as well as the fungal/bacterial SSU ratio were
303 part of one interconnected node hub. More specifically, all ARGs showed a direct or indirect
304 connection (one common link) to the fungal/bacterial SSU ratio. Furthermore, all extracted
305 correlations directly or indirectly connecting ARGs with fungal/bacterial SSU ratio were positive
306 (Fig. 4). Specifically, *bla*$_{TEM}$ was directly positively correlated with fungal/bacterial SSU ratio and
307 further indirectly connected through the bacterial order of *Corynebacteriales*, supporting the
308 previously detected strong correlation of this ARG. Moreover, the relative abundance of *bla*$_{TEM}$
309 was the only explanatory factor connected to the aminoglycoside ARG *aph*(3') (Fig. 4), which
310 supports the previously indicated weaker positive correlation with fungal relative abundance.
311 Meanwhile, *bla*$_{OXA}$ was exclusively connected to the bacterial order of *Rhizobiales*, which was in
312 turn providing the indirect link through positive correlation with the fungal/bacterial SSU ratio
313 based on Spearman correlations.

314 In addition to the ARGs, fungal abundance was positively correlated with a number of individual
315 bacterial taxa, however not a single antagonistic interaction was observed (Fig. 4). These
316 observed positive correlations indicate potential mutualistic interactions. Selection for specific
317 bacterial taxa was driven by their ability to co-exist with fungi, despite the potential production of
318 secondary metabolites by fungi with negative effects on bacterial growth. Since β-lactam ARGs
319 confer resistance to β-lactam antibiotic, which are commonly produced antibiotics by fungi (Aly
320 et al., 2011), we hypothesize that these ARGs could potentially have enabled the co-existence

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

of several of these bacterial taxa with fungi, hence promoting their co-occurrence as individual, interconnected nodes within the correlation network centering around fungal abundance.

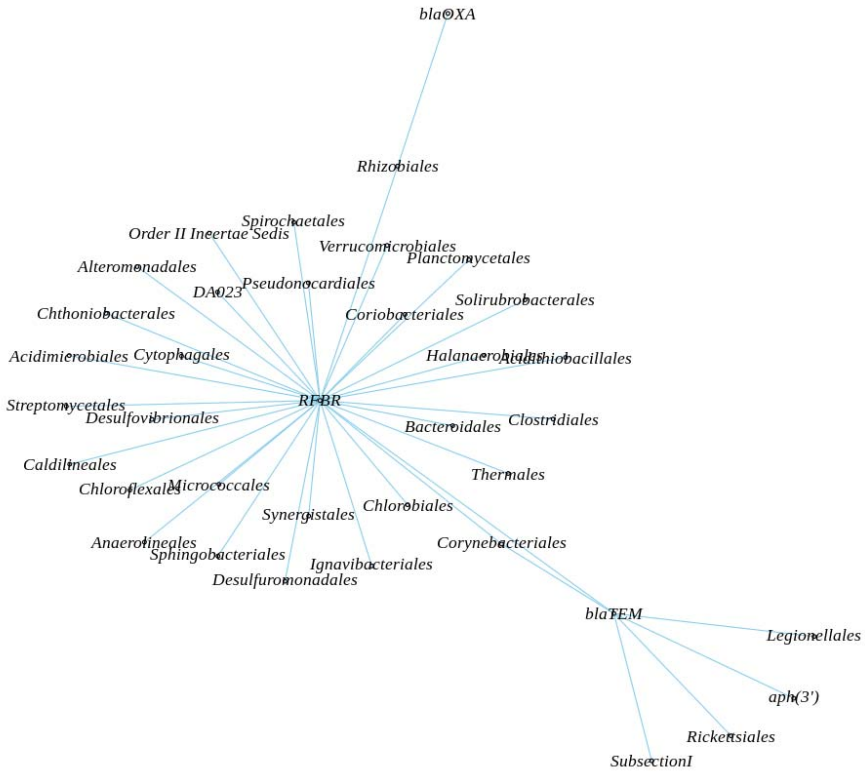


Figure 4. Extract of correlations including either the ribosomal fungal/bacterial SSU ratio (RFBR) or any of the detected ARGs from the full co-occurrence network (Fig. S3). Only significant correlations based on Spearman Correlation with Benjamini-Hochberg correction for multiple testing (p<0.05) are shown. All extracted correlations within this network were significantly positive.

In summary, ARGs were regularly only directly connected to a minor proportion of taxa but rather directly or indirectly connected to fungal abundance with positive correlations. Consequently, fungal abundance might serve as a better indicator for the abundance of certain ARGs in GW microbiomes than the bacterial community composition itself.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

3.5 Summary of results

In the present *in silico* study we identified the common and core ARGs that make up the global GW resistome and elucidated potential drivers underlying their abundance patterns. The common GW ARGs conferred resistance to 13 antibiotic classes, while the core resistome was made up of six ARGs conferring resistance to sulfonamides, β-lactams and aminoglycosides. Local patterns regarding the intensity of anthropogenic factors were identified as a driving force behind the distribution of ARGs conferring resistance to the synthetic antibiotic class of sulfonamides. However, for β-lactams - natural, fungal-derived antibiotics (Nesme and Simonet, 2015) - the relative abundance of these fungi provided a main explanatory variable. A previous investigation on global soil microbiota supports such a correlation of fungi with total ARGs (Bahram et al., 2018). While across soils co-selective effects for a number of antibiotic classes could be detected, in GW metagenomes only the β-lactam ARGs bla_{TEM} and bla_{OXA} were directly correlated with fungal abundance. In addition, fungal abundance served as indirect indication for the aminoglycoside ARG *aph*(3'), which belonged to the core GW resistome and indirectly correlated with fungal abundance through co-occurrence network analysis.

4. Conclusion

Overall we show that the re-analysis of publicly available data is a valuable tool for testing hypotheses currently present in the microbial ecology spectrum and elucidating potential global relationships between different microbial groups in GW environments. Specifically, we demonstrated that in the pristine GW environments, the global resistome is dominated by a small number of ARGs and that their abundance profiles, where mostly influenced by local conditions, while they can be partially shaped by microbe-microbe interactions. By using these *in silico* approaches we can pinpoint and identify potential microbe-microbe interactions for further verification in controlled laboratory experiments. In addition, we demonstrated that the bacterial/fungal SSU ratio could act as a direct and indirect indicator for the abundance of

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

357 specific ARGs in GW environments. We expect that with the increase of publicly available data,
358 such *in silico* meta-analyses will be able to further identify ecological interactions in
359 understudied environments in the future.

360 5. Conflict of interest

361 The authors declare no conflict of interest.

362 6. Acknowledgements

363 We deeply thank the researchers that uploaded and provided their sequencing data in the NCBI
364 SRA databases, as without their contribution this study would have been impossible.

365 7. Funding

366 This work was supported by the JPI AMR - EMBARK and the ANTIVERSA project funded by
367 the Bundesministerium für Bildung, und Forschung under grant numbers & F01KI1909A &
368 01LC1904A, and the European Union's Horizon 2020 research and innovation program under
369 the PRIMA program supported by the European Union under grant agreement No 1822. JBP
370 acknowledges funding from the Swedish Research Council (VR; grant 2019-00299) under the
371 frame of JPI AMR (EMBARK; JPIAMR2019-109) and the Data-Driven Life Science (DDLS)
372 program supported by the Knut and Alice Wallenberg Foundation (KAW 2020.0239).
373 Responsibility for the information and views expressed therein lies entirely with the authors.

374 8. Data availability

375 Data and the R and shell scripts for the workflow of analysis have been uploaded in
376 https://github.com/JonKampouris/GW_Resistome.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

9. References

1) Aly, A.H., Debbab, A., Proksch, P., 2011. Fungal endophytes: Unique plant inhabitants with great promises. Appl. Microbiol. Biotechnol. 90, 1829–1845. https://doi.org/10.1007/s00253-011-3270-y

2) Alygizakis, N.A., Urík, J., Beretsou, V.G., Kampouris, I., Galani, A., Oswaldova, M., Berendonk, T., Oswald, P., Thomaidis, N.S., Slobodnik, J., Vrana, B., Fatta-Kassinos, D., 2020. Evaluation of chemical and biological contaminants of emerging concern in treated wastewater intended for agricultural reuse. Environ. Int. 138, 105597. https://doi.org/10.1016/j.envint.2020.105597

3) Avisar, D., Lester, Y., Ronen, D., 2009. Sulfamethoxazole contamination of a deep phreatic aquifer. Sci. Total Environ. 407, 4278–4282. https://doi.org/10.1016/j.scitotenv.2009.03.032

4) Bahram, M., Hildebrand, F., Forslund, S.K., Anderson, J.L., Soudzilovskaia, N.A., Bodegom, P.M., Bengtsson-Palme, J., Anslan, S., Coelho, L.P., Harend, H., Huerta-Cepas, J., Medema, M.H., Maltz, M.R., Mundra, S., Olsson, P.A., Pent, M., Põlme, S., Sunagawa, S., Ryberg, M., Tedersoo, L., Bork, P., 2018. Structure and function of the global topsoil microbiome. Nature 560, 233–237. https://doi.org/10.1038/s41586-018-0386-6

5) Bengtsson-Palme, J., Larsson, D.G.J., 2016. Concentrations of antibiotics predicted to select for resistant bacteria: Proposed limits for environmental regulation. Environ. Int. 86, 140–149. https://doi.org/10.1016/j.envint.2015.10.015

6) Bortolaia, V., Kaas, R.S., Ruppe, E., Roberts, M.C., Schwarz, S., Cattoir, V., Philippon, A., Allesoe, R.L., Rebelo, A.R., Florensa, A.F., Fagelhauer, L., Chakraborty, T.,

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Neumann, B., Werner, G., Bender, J.K., Stingl, K., Nguyen, M., Coppens, J., Xavier, B.B., Malhotra-Kumar, S., Westh, H., Pinholt, M., Anjum, M.F., Duggett, N.A., Kempf, I., Nykäsenoja, S., Olkkola, S., Wieczorek, K., Amaro, A., Clemente, L., Mossong, J., Losch, S., Ragimbeau, C., Lund, O., Aarestrup, F.M., 2020. ResFinder 4.0 for predictions of phenotypes from genotypes. J. Antimicrob. Chemother. 75, 3491–3500. https://doi.org/10.1093/jac/dkaa345.

7) Buchfink, B., Xie, C., Huson, D.H., 2014. Fast and sensitive protein alignment using DIAMOND. Nat. Methods 12, 59–60. https://doi.org/10.1038/nmeth.3176

8) Cacace, D., Fatta-Kassinos, D., Manaia, C.M., Cytryn, E., Kreuzinger, N., Rizzo, L., Karaolia, P., Schwartz, T., Alexander, J., Merlin, C., Garelick, H., Schmitt, H., de Vries, D., Schwermer, C.U., Meric, S., Ozkal, C.B., Pons, M.N., Kneis, D., Berendonk, T.U., 2019. Antibiotic resistance genes in treated wastewater and in the receiving water bodies: A pan-European survey of urban settings. Water Res. 162, 320–330. https://doi.org/10.1016/j.watres.2019.06.039

9) Caucci, S., Karkman, A., Cacace, D., Rybicki, M., Timpel, P., Voolaid, V., Gurke, R., Virta, M., Berendonk, T.U., 2016. Seasonality of antibiotic prescriptions for outpatients and resistance genes in sewers and wastewater treatment plant outflow. FEMS Microbiol. Ecol. 92, fiw060. https://doi.org/10.1093/femsec/fiw060

10) Cock, P.J.A., Antao, T., Chang, J.T., Chapman, B.A., Cox, C.J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., De Hoon, M.J.L., 2009. Biopython: Freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics 25, 1422–1423. https://doi.org/10.1093/bioinformatics/btp163

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

11) Coelho, L.P., Alves, R., Monteiro, P., Huerta-Cepas, J., Freitas, A.T., Bork, P., 2019. NG-meta-profiler: Fast processing of metagenomes using NGLess, a domain-specific language. Microbiome 7, 1–10. https://doi.org/10.1186/s40168-019-0684-8

12) Csardi, Gabor, and Tamas Nepusz. 2005. The Igraph Software Package for Complex Network Research. InterJournal Complex Systems, 1695

13) Flynn, T.M., Sanford, R.A., Ryu, H., Bethke, C.M., Levine, A.D., 2013. Functional microbial diversity explains groundwater chemistry in a pristine aquifer Functional microbial diversity explains groundwater chemistry in a pristine aquifer. BMC Microbiol. 13, 1.

14) Gatica, J., Jurkevitch, E., Cytryn, E., 2019. Comparative metagenomics and network analyses provide novel insights into the scope and distribution of β-lactamase homologs in the environment. Front. Microbiol. 10, 1–11. https://doi.org/10.3389/fmicb.2019.00146

15) Gatica, J., Yang, K., Pagaling, E., Jurkevitch, E., Yan, T., Cytryn, E., 2015. Resistance of undisturbed soil microbiomes to ceftriaxone indicates extended spectrum β-lactamase activity. Front. Microbiol. 6, 1–11. https://doi.org/10.3389/fmicb.2015.01233

16) Griebler, C., Avramov, M., 2015. Groundwater ecosystem services : a review 34, 355–367. https://doi.org/10.1086/679903.

17) Griebler, C., Lueders, T., 2009. Microbial biodiversity in groundwater ecosystems. Freshw. Biol. 54, 649–677. https://doi.org/10.1111/j.1365-2427.2008.02013.x

18) Hernando-Amado, S., Coque, T.M., Baquero, F., Martínez, J.L., 2019. Defining and combating antibiotic resistance from One Health and Global Health perspectives. Nat. Microbiol. 4, 1432–1442. https://doi.org/10.1038/s41564-019-0503-9

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

19) Herrmann, M., Wegner, C.E., Taubert, M., Geesink, P., Lehmann, K., Yan, L., Lehmann, R., Totsche, K.U., Küsel, K., 2019. Predominance of Cand. Patescibacteria in groundwater is caused by their preferential mobilization from soils and flourishing under oligotrophic conditions. Front. Microbiol. 10, 1–15. https://doi.org/10.3389/fmicb.2019.01407

20) Johnston, S.R., Boddy, L., Weightman, A.J., 2016. Bacteria in decomposing wood and their interactions with wood-decay fungi. FEMS Microbiol. Ecol. 92, 1–12. https://doi.org/10.1093/femsec/fiw179

21) Jones, E.R., Van Vliet, M.T.H., Qadir, M., Bierkens, M.F.P., 2021. Country-level and gridded estimates of wastewater production, collection, treatment and reuse. Earth Syst. Sci. Data 13, 237–254. https://doi.org/10.5194/essd-13-237-2021

22) Kampouris, I.D., Agrawal, S., Orschler, L., Cacace, D., Kunze, S., Berendonk, T.U., Klümper, U., 2021. Antibiotic resistance gene load and irrigation intensity determine the impact of wastewater irrigation on antimicrobial resistance in the soil microbiome. Water Res. 193. https://doi.org/10.1016/j.watres.2021.116818

23) Kampouris, I.D., Alygizakis, N., Klümper, U., Agrawal, S., Lackner, S., Cacace, D., Kunze, S., Thomaidis, N.S., Slobdonik, J., Berendonk, T.U., 2022. Elevated levels of antibiotic resistance in groundwater during treated wastewater irrigation associated with infiltration and accumulation of antibiotic residues. J. Hazard. Mater. 423, 127155. https://doi.org/10.1016/j.jhazmat.2021.127155

24) Karkman, A., Pärnänen, K., Larsson, D.G.J., 2019. Fecal pollution can explain antibiotic resistance gene abundances in anthropogenically impacted environments. Nat. Commun. 10, 1–8. https://doi.org/10.1038/s41467-018-07992-3

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

25) Kassambara, A., 2019. ggpubr R Package: ggplot2-Based Publication Ready Plots R package version 0.2.3 https://CRAN.R-project.org/package=ggpubr.

26) Laxminarayan, R., Duse, A., Wattal, C., Zaidi, A.K.M., Wertheim, H.F.L., Sumpradit, N., Vlieghe, E., Hara, G.L., Gould, I.M., Goossens, H., Greko, C., So, A.D., Bigdeli, M., Tomson, G., Woodhouse, W., Ombaka, E., Peralta, A.Q., Qamar, F.N., Mir, F., Kariuki, S., Bhutta, Z.A., Coates, A., Bergstrom, R., Wright, G.D., Brown, E.D., Cars, O., 2013. Antibiotic resistance-the need for global solutions. Lancet Infect. Dis. 13, 1057–1098. https://doi.org/10.1016/S1473-3099(13)70318-9

27) Li, H., 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM 00, 1–3.

28) Li, H., Durbin, R., 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26, 589–595. https://doi.org/10.1093/bioinformatics/btp698

29) Liao, Z., Chen, Z., Xu, A., Gao, Q., Song, K., Liu, J., Hu, H.Y., 2021. Wastewater treatment and reuse situations and influential factors in major Asian countries. J. Environ. Manage. 282. https://doi.org/10.1016/j.jenvman.2021.111976

30) Linares, J.F., Gustafsson, I., Baquero, F., Martinez, J.L., 2006. Antibiotics as intermicrobiol signaling agents instead of weapons. Proc. Natl. Acad. Sci. U. S. A. 103, 19484–19489. https://doi.org/10.1073/pnas.0608949103

31) Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal, 17(1), pp. 10-12. doi:https://doi.org/10.14806/ej.17.1.200

32) Miao, V, Davies, J. Actinobacteria: The good, the bad, and the ugly. Antonie van Leeuwenhoek, Int J Gen Mol Microbiol 2010; 98: 143–150.

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

33) Nawaz, A., Purahong, W., Lehmann, R., Herrmann, M., Totsche, K.U., Küsel, K., Wubet, T., Buscot, F., 2018. First insights into the living groundwater mycobiome of the terrestrial biogeosphere. Water Res. 145, 50–61. https://doi.org/10.1016/j.watres.2018.07.067

34) Nesme J, Simonet P. The soil resistome: A critical review on antibiotic resistance origins, ecology and dissemination potential in telluric bacteria. Environ Microbiol 2015; 17: 913–930.

35) Oksanen; J., Blanchet; F.G., Friendly; M., Kindt; R., Legendre; P., McGlinn; D., Minchin; P.R., O'Hara; R.B., Simpson; G.L., Solymos; P.M. Stevens; H.H., Szoecs; E., Wagner H., 2019. vegan R package: an R package for community ecologists. R package version 2.5-6. https://cran.rproject.org/package=vegan

36) Pedersen T (2022). ggraph: An Implementation of Grammar of Graphics for Graphs and Networks. https://ggraph.data-imaginist.com, https://github.com/thomasp85/ggraph.

37) R Core Team, R: A language and environment for statistical computing R Foundation for Statistical Computing, Vienna, Austria (2019)

38) Retter, A., Karwautz, C., Griebler, C., 2021. Groundwater microbial communities in times of climate change. Curr. Issues Mol. Biol. 41, 509–538. https://doi.org/10.21775/cimb.041.509

39) Sonthiphand, P., Ruangroengkulrith, S., Mhuantong, W., Charoensawan, V., Chotpantarat, S., Boonkaewwan, S., 2019. Metagenomic insights into microbial diversity in a groundwater basin impacted by a variety of anthropogenic activities. Environ. Sci. Pollut. Res. 26, 26765–26781. https://doi.org/10.1007/s11356-019-05905-5

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

514 40) Spielmeyer A, Höper H, Hamscher G. Long-term monitoring of sulfonamide leaching
515 from manure amended soil into groundwater. Chemosphere 2017; 177: 232–238.

516 41) Szekeres, E., Chiriac, C.M., Baricz, A., Szőke-Nagy, T., Lung, I., Soran, M.L., Rudi, K.,
517 Dragos, N., Coman, C., 2018. Investigating antibiotics, antibiotic resistance genes, and
518 microbial contaminants in groundwater in relation to the proximity of urban areas.
519 Environ. Pollut. 236, 734–744. https://doi.org/10.1016/j.envpol.2018.01.107

520 42) Taş, N., Brandt, B.W., Braster, M., van Breukelen, B.M., Röling, W.F.M., 2018.
521 Subsurface landfill leachate contamination affects microbial metabolic potential and
522 gene expression in the Banisveld aquifer. FEMS Microbiol. Ecol. 94, 1–12.
523 https://doi.org/10.1093/femsec/fiy156

524 43) Vaz-Moreira, I., Nunes, O.C., Manaia, C.M., 2014. Bacterial diversity and antibiotic
525 resistance in water habitats: Searching the links with the human microbiome. FEMS
526 Microbiol. Rev. 38, 761–778. https://doi.org/10.1111/1574-6976.12062

527 44) Wang, H., Su, X., Su, J., Zhu, Y., Ding, K., 2022. Profiling the antibiotic resistome in soils
528 between pristine and human-affected sites on the Tibetan Plateau. J. Environ. Sci.
529 (China) 111, 442–451. https://doi.org/10.1016/j.jes.2021.04.019

530 45) Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer (2016)

531 46) Wickham, H., François, R., Henry, L., Müller, K. 2022. dplyr: A Grammar of Data
532 Manipulation. https://dplyr.tidyverse.org, https://github.com/tidyverse/dplyr.

533 47) Wickham; H., Averick; M., Bryan; J., Chang; W., McGowan; L., François; R., Grolemund;
534 G., Hayes; A., Henry; L., Hester; J., Kuhn; M., Pedersen; T., Miller; E., Bache; S.,
535 Müller; K., Ooms; J., Robinson; D., Seidel; D., Spinu; V., Takahashi; K., Vaughan; D.,

bioRxiv preprint doi: https://doi.org/10.1101/2022.11.14.516424; this version posted November 15, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

536 Wilke; C., Woo; K., Yutani; H. Welcome to the Tidyverse. J. Open Source Softw. 2019;
537 4 (43); 1686. https://doi.org/10.21105/joss.01686.

538 48) Yan L, Herrmann M, Kampe B, Lehmann R, Totsche KU, Küsel K. Environmental
539 selection shapes the formation of near-surface groundwater microbiomes. Water Res
540 2020; 170: 115341.

541 49) Zaouri, N., Jumat, M.R., Cheema, T., Hong, P.Y., 2020. Metagenomics-based evaluation
542 of groundwater microbial profiles in response to treated wastewater discharge.
543 Environ. Res. 180, 108835. https://doi.org/10.1016/j.envres.2019.108835

544 50) Zhang, Y. Zhang, Yuanzhu, Kuang, Z., Xu, J., Li, C., Li, Y., Jiang, Y., Xie, J., 2019.
545 Comparison of Microbiomes and Resistomes in Two Karst Groundwater Sites in
546 Chongqing, China. Groundwater 57, 807–818. https://doi.org/10.1111/gwat.12924