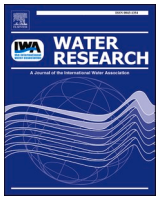
Water Research 281 (2025) 123553

Contents lists available at ScienceDirect

Water Research

journal homepage: www.elsevier.com/locate/watres



Anthropogenic imprint on riverine plasmidome diversity and proliferation of antibiotic resistance genes following pollution and urbanization

Kenia Barrantes-Jiménez [a,b], Franck Lejzerowicz [c], Tam Tran [d], Melany Calderón-Osorno [e], Luis Rivera-Montero [b], César Rodríguez-Sánchez [f], Odd-Gunnar Wikmark [d], Alexander Eiler [c], Hans-Peter Grossart [g,h], María Arias-Andrés [i,*], Keilor Rojas-Jiménez [j,*]

[a] Doctorado en Ciencias Naturales para el Desarrollo (DOCINADE), Instituto Tecnológico de Costa Rica, Universidad Nacional and Universidad Estatal a Distancia, San José, Costa Rica
[b] Health Research Institute, University of Costa Rica, P.O. Box: 11501-2060, San José, Costa Rica
[c] Section for Aquatic Biology and Toxicology, Blindernveien 31 0371 Oslo, University of Oslo, Norway
[d] NORCE, Siva Innovasjonssenter, Sykehusvn 21, 9019 Tromsø, Norway
[e] Costa Rica National High Technology Center (CeNAT), P.O. Box: 1174-1200, San José, Costa Rica
[f] Faculty of Microbiology & Research Center for Tropical Diseases (CIET), University of Costa Rica, P.O. Box: 11501-2060, San José, Costa Rica
[g] Leibniz Institute for Freshwater Ecology and Inland Fisheries, IGB, Department 3, Plankton and Microbial Ecology, Zur Alten Fischerhuette 2, 16775 Stechlin, Germany
[h] Institute of Biochemistry and Biology, Potsdam University, Maulbeerallee 2, D-14469 Potsdam, Germany
[i] Instituto Regional de Estudios en Sustancias Tóxicas (IRET), Universidad Nacional, Campus Omar Dengo, P.O. Box 86-3000, Heredia, Costa Rica
[j] Biology School, University of Costa Rica, P.O. Box: 11501-2060, San José, Costa Rica

ARTICLE INFO

Keywords:
Plasmids
Antibiotic resistance genes (ARGs)
Tropical rivers
Anthropogenic pollution
Microbial ecology and evolution
Microbial diversity
Cargo genes enrichment
Health risk

ABSTRACT

Plasmids are key determinants in microbial ecology and evolution, facilitating the dissemination of adaptive traits and antibiotic resistance genes (ARGs). Although the molecular mechanisms governing plasmid replication, maintenance, and transfer have been extensively studied, the specific impacts of urbanization-induced pollution on plasmid ecology, diversity, and associated ARGs in tropical regions remain underexplored. This study investigates these dynamics in a tropical aquatic ecosystem, providing novel insights into how pollution shapes plasmid composition and function. In contrast to the observed decrease in chromosomal diversity, we demonstrate that pollution associated with urbanization increases the diversity and taxonomic composition of plasmids within a bacterial community (plasmidome). We analyzed eighteen water and sediment metagenomes, capturing a gradient of pollution and ARG contamination along a tropical urban river. Plasmid and chromosomal diversity profiles were found to be anti-correlated. Plasmid species enrichment along the pollution gradient led to significant compositional differences in water samples, where differentially abundant species suggest plasmid maintenance within specific taxonomic classes. Additionally, the diversity and abundance of ARGs related to the plasmidome increased concomitantly with the intensity of fecal and chemical pollution. These findings highlight the critical need for targeted plasmidome studies to better understand plasmids' environmental spread, as their dynamics are independent of chromosomal patterns. This research is crucial for understanding the consequences of bacterial evolution, particularly in the context of environmental and public health.

1. Introduction

Plasmids play a crucial role in the evolution of bacteria by facilitating the dissemination of genes that confer adaptive advantages (Partridge et al., 2018; Rodríguez-Beltrán et al., 2021). Unlike chromosomal genes, plasmids are often maintained at dynamically adjusted copies per cell. This dynamic polyploidy leads to plasmid-encoded genes being governed by different evolutionary dynamics than the chromosomal genes, including more rapid evolution (due to higher copy numbers), genetic dominance, heteroplasmy (the presence of genetic variants within the same cell), and segregational drift (Rodríguez-Beltrán et al., 2021). This distinct evolution is attributed to the dynamic nature of plasmids in

* Corresponding authors.
E-mail addresses: maria.arias.andres@una.ac.cr (M. Arias-Andrés), keilor.rojas@ucr.ac.cr (K. Rojas-Jiménez).

https://doi.org/10.1016/j.watres.2025.123553
Received 24 October 2024; Received in revised form 22 March 2025; Accepted 24 March 2025
Available online 25 March 2025
0043-1354/© 2025 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

bacterial populations, leading to high levels of genetic plasticity, enabling populations to adapt swiftly to environmental changes and selective pressures. Consequently, all plasmids in a bacterial community (plasmidome) may exhibit genetic and ecological patterns distinct from the host chromosomes, influencing bacterial survival and competitiveness in diverse environments (Rodríguez-Beltrán et al., 2021; Slater et al., 2008; Smillie et al., 2010).

While carrying plasmids can sometimes reduce bacterial fitness, they often provide advantages to the host (Larsson and Flach, 2022; Newbury et al., 2022; Topp et al., 2018). For example, they provide resistance against environmental challenges, including anthropogenic stressors such as antibiotics. Antibiotic resistance genes (ARGs) are often associated with a diverse array of mobile genetic elements (MGEs), enabling these genes to spread to new hosts of different taxonomic groups through horizontal gene transfer (HGT) (Arias-Andres et al., 2018; Gillings 2017). However, despite the significance of plasmids in microbial ecology and evolution, the understanding of their composition, structure, and diversity patterns in the environment, especially in aquatic ecosystems, remains limited (Ebmeyer et al., 2021; Rodríguez-Beltrán et al., 2021; Smalla et al., 2018).

Increasing human activities and urbanization significantly impact aquatic ecosystems, introducing pharmaceuticals, disinfectants, personal care products, and other anthropogenic stressors that also provide such mentioned selective pressures for microbial HGT dynamics (Ramírez-Morales et al., 2021; Venegas González et al., 2023; Wilkinson et al., 2022). This affects the dissemination of ARGs, among other traits often associated with a diverse array of MGEs (Ebmeyer et al., 2021; Smalla et al., 2018). Plasmids play a significant role in providing resistance against antibiotics (Arias-Andres et al., 2018; Gillings, 2017). However, as the understanding of plasmid composition, structure, and diversity patterns in the environment remains limited, so is our interpretation of antibiotic resistance dynamics in aquatic ecosystems (Ebmeyer et al., 2021; Rodríguez-Beltrán et al., 2021; Smalla et al., 2018). Thus, examining the plasmidome, in aquatic environments under anthropogenic stress is crucial, particularly in developing countries where the management of antibiotics and waste disposal are often inadequate (Larsson and Flach, 2022; Suzuki et al., 2017). Beyond understanding plasmid dynamics and ecology, it also sheds light on potential risks to human health and the environment posed by the increasing dissemination of antimicrobial resistance amplified by these MGEs (Bottery, 2022; Jiang et al., 2022; Larsson and Flach, 2022).

Recent advancements in metagenomics, such as the development of machine-learning models, have significantly improved our ability to identify and characterize plasmids within complex microbial ecosystems. These tools have uncovered diverse plasmid systems within human gut metagenomes, revealing various genes that enhance microbial fitness across different conditions (Yu et al., 2024). Such studies underscore the critical role plasmids play in microbial ecology by facilitating HGT, which results in adaptive advantages, including antibiotic resistance. However, many existing tools for plasmid analysis have been developed on an "*ad hoc* basis." Therefore, creating more comprehensive bioinformatics pipelines that integrate these tools and allow for the thorough analysis of entire plasmidomes, regardless of their chromosomal context, remains essential (Stockdale et al., 2022).

Few studies separated plasmids from chromosomes to study microbiomes or assessed the impact of environmental pollution on their diversity and the ARGs they carry (Shintani et al., 2015; van Passel et al., 2006). Plasmidomes are often more conserved than the corresponding chromosomal microbiomes, emphasizing their unique ecological role in maintaining latent genetic functionalities across diverse environments (Kothari et al., 2019). Similarly, it has been highlighted that plasmid diversity can act as a biomarker for health and disease states, further underscoring the importance of distinguishing plasmids in metagenomic analyses to uncover their functional contributions (Stockdale et al., 2022). This distinction enables a more accurate understanding of microbial community dynamics and plasmids' ecological and evolutionary roles, particularly in environments impacted by human activity (Stockdale and Hill, 2023).

Here, we analyze the plasmidome of a highly polluted tropical river severely impacted by urbanization. This demonstrates that differentiating plasmids from chromosomal populations may have important implications in microbial ecology and evolution. Our thorough analyses of sediment and water column samples collected along a pollution gradient in a Costa Rican River based on 18 metagenomes confirm that acute urban pollution affects the diversity and genetic composition of the bacterial plasmidome. By selecting plasmids carrying ARGs, we hypothesized that urban pollution increases the plasmidome richness and diversity. We propose that plasmids and some genes within them may serve as unique biomarkers and, thus, indicators of urban pollution and the risk of ARGs spreading in the environment.

2. Material and methods

2.1. Site description and sampling

This study was conducted along a gradient in pollution and anthropogenic activities (urbanization). Samples from the water column and sediment were collected at three sites within the main course of the Virilla River, located in the Western Central Valley of Costa Rica (Fig. 1). The Virilla River watershed is part of the larger Grande de Tarcoles River basin, which drains into the Pacific Ocean. Urbanization and changing land use patterns within the watershed have increased pollution levels, particularly from untreated municipal sewage and industrial effluents (Herrera-Murillo et al., 2019; Mena-Rivera et al., 2018). Notably, Site 3, located downstream of the San José Wastewater Treatment Plant (SJ-WWTP), exhibits the highest pollution levels, including antibiotics, making it a hotspot for ARGs dissemination and plasmid transfer (Mendoza-Guido et al., 2024). Consistent with this, the Virilla River study demonstrated that areas with urban and grazing land use exhibit higher prevalence and diversity of resistance and virulence genes in *Escherichia coli*. Resistance genes, such as bla_{TEM}, were frequently detected and associated with anthropogenic activities, including wastewater discharge and livestock management. This highlights the compounded impact of urbanization and environmental dynamics on microbial pollution and associated public health risks (Morales-Mora et al., 2025).

The sampling was carried out during the dry season in April 2022. At each site, triplicate samples were taken, resulting in 9 samples per sample type and 18 samples in total. Site 1, the highest elevation site, is at an altitude of 2020 m a.s.l. (9°59′9″N, 83°56′35″W). The primary land uses in this area are cattle raising and plant cultivation. Site 2, the intermediate elevation site, is 1185 m a.s.l. (9°57′47″N, 84°5′51″W). This site encompasses residential, commercial, industrial, and urban land uses. Site 3, the lowest elevation site, is 720 m a.s.l. (9°56′46″N, 84°13′22″W). The land use is predominantly urban and industrial, and Site 3 is downstream of the San José Wastewater Treatment Plant (SJ-WWTP). The SJ-WWTP facility is the largest in the country and is designed to serve over one million people in the Greater Metropolitan Area of San José. The plant primarily employs physical and chemical treatment processes to significantly reduce pollutants such as total suspended solids (TSS), biological oxygen demand (BOD), and oils from the wastewater entering the rivers María Aguilar, Tiribí, and Torres, all of them part of the Virilla River watershed (Mora-Aparicio et al., 2022).

Freshwater and sediment samples were collected to provide diverse insights into the pollution levels within riverine habitats. Due to their ability to absorb and retain substances, sediments accumulate ARGs and antimicrobials, making them valuable for studying the long-term impact of pollution and the evolution of AMR. In contrast, water samples can reveal the present circulation of ARGs and antimicrobials, indicating ongoing pollution sources. Also, analyzing water columns helps to monitor the effectiveness of pollution control measures, such as WWTPs, and identify emerging health threats. Therefore, the combined analysis

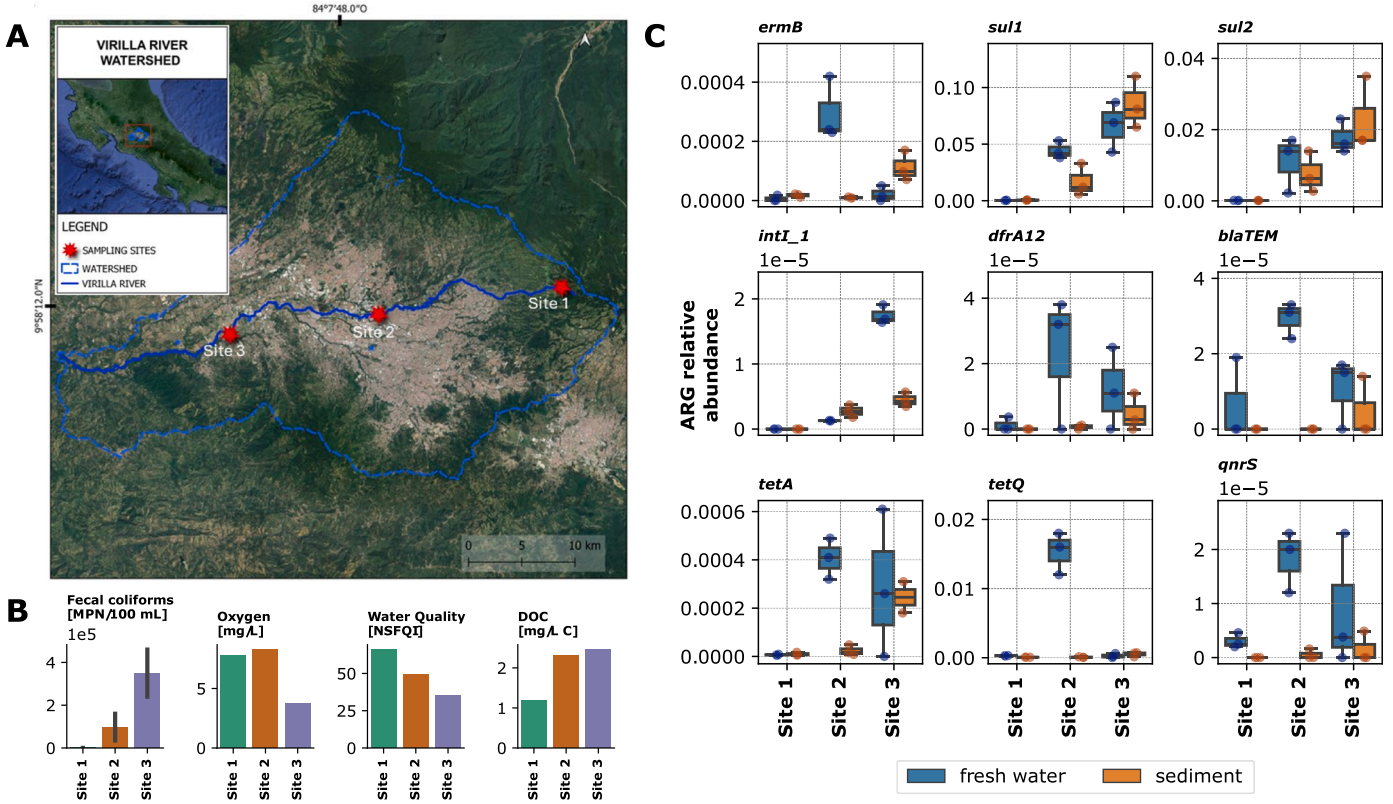


Fig. 1. Chemical pollution and biological pollution gradient along the Virilla River. (A) Virilla River watershed and sampling sites (a wastewater treatment plant is located between sites 2 and 3). (B) Quality indicators are measured in the water column at each site, incl. the National Sanitation Foundation Quality Index (NSFQI). (C) Relative abundances of 9 ARGs were measured by qPCR in water and sediment samples at each site.

of riverine freshwater and sediments offers a comprehensive understanding of pollution dynamics. Water samples were collected using sterile containers and gloves. Four glass bottles were rinsed with surface water from the river three times for water column samples. Then, one liter of water was sampled in each bottle. Sediments were collected using a sterile spoon and plastic bags from the river's shore under running water at a depth of approximately 30 cm. Samples were placed into a sterile bag and sealed within another sterile bag. All samples, water, and sediment were transported to different laboratories under cold conditions and analyzed within 20 h.

2.2. Microbiological and chemical analysis

Microbiological and chemical indicators were measured only in water column samples. Fecal coliforms, *E. coli*, and *Enterococcus faecalis (E.faecalis)* were quantified using the most probable number technique (standard methods 9221E, 9221F, and 9230B) (American Public Health Association et al., 2017). Oxygen saturation, conductivity, pH, and temperature were analyzed using a multiparameter probe (YSI, model 85 A, Yellow Springs, OH, USA).

Dissolved Organic Carbon (DOC) analysis (Table S1) was performed in the Environmental Pollution Research Center (CICA) laboratory at UCR using the standard method 5310A (American Public Health Association et al., 2017). Turbidity analysis was conducted in the Health Research Institute (INISA) laboratory (Table S1) using the standard method 2510B (American Public Health Association et al., 2017). Analyses of residues of anthropogenic contaminants such as caffeine and terbutryn (Table S1) were performed in the Central American Institute for Studies on Toxic Substances (IRET) Laboratory at UNA. Samples analyzed by gas chromatography (GC) were extracted by solid phase extraction (SPE) and processed using GC coupled to mass spectrometry (GC–MS) and electron capture detection (GCECD) in an Agilent, 7890A-5975C equipment.

Also, the National Sanitation Foundation Water Quality Index (NSFQI) was calculated according to references (Brown et al., 1972; Marselina et al., 2022; Uddin et al., 2021). Five parameters were included in the calculation: fecal coliforms, pH, temperature, turbidity, and oxygen saturation. By considering these parameters, the NSFQI provides a comprehensive overview of water quality, allowing the identification of potential anthropogenic pollution and classifying surface waters into five categories: very good, good, fair, poor, and very poor (Ichwana et al., 2016; Marselina et al., 2022).

2.3. Environmental DNA extraction and quantification

Water samples were prefiltered via a filtration device (Sartorius®, Göttingen, Germany) equipped with an 80-μm glass fiber prefilter (13,400–47-Q, Sartorious®, Germany) to remove larger particles. The prefiltered samples were then filtered onto a 0.22-μm cellulose nitrate filter (47-mm diameter;11,327–47–N, Sartorius®, Germany).

According to the kit's directions, DNA from filters and sediments was extracted with the DNeasy PowerSoil Pro (Qiagen, Venlo, The Netherlands). Two hundred fifty milligrams of the sample were used for sediments, and one filter (equivalent to 250 ml filtration volume) was extracted for the water sample. The extracts' DNA concentration (and purity) was measured using a NanoDrop 2000 spectrophotometer (ThermoFisher Scientific, USA). It ranged between 7.5 and 50 ng/μl for the water samples (A260/280 values: 1.6–2.1), and between 7.0 and 59 ng/μl (A260/280 values: 1.6–1.9) for the sediment samples. These extracts were aliquoted and stored at −80 °C before ARGs quantification and sequencing.

2.4. Quantification of antimicrobial resistance genes

Standard quantitative polymerase chain reaction (qPCR) using SybrGreen chemistry was performed to analyze ARGs in DNA samples from the water column and sediments. ARG primer sequences for *16S rRNA, intI-1, sul-1, sul-2, tet(A)*, *tet(Q), dfrA12, qnrS, ermB,* bla_{TEM}, and $bla_{CTX\text{-}M}$, as well as reaction conditions, primer concentration, expected amplicon, and reaction efficiency.

Table S2 details the critical parameters for qPCR experiments targeting 11 ARGs, including primer sequences, expected product sizes (85–190 bp), primer concentrations (200–600 nM), and annealing temperatures (79.8–90.02 °C). The amplification protocol involved initial denaturation at 95 °C for 2–10 min, followed by 40 cycles of denaturation (95 °C for 3–5 s) and annealing/extension (60 °C for 30 s). Reaction efficiencies ranged from 73.5 % to 104.09 %, with R^2 values consistently above 0.993, ensuring reliable and reproducible amplification.

These genes are associated with resistance to eight different classes of antibiotics, namely sulfonamides (*sul-1* and *sul-2*), trimethoprim (*dfrA12*), beta-lactams (bla_{TEM}), quinolones (*qnrS*), tetracyclines (*tet(A)* and *tet(Q)*), and macrolides, lincosamines and streptogramins (*ermB*). They were selected based on a review of scientific literature highlighting the presence of antibiotics such as sulfonamides, tetracycline derivatives, trimethoprim, and beta-lactams as common contaminants in surface waters in the country (Spongberg et al., 2011; De la Cruz et al., 2014; Ramírez-Morales et al., 2021). This selection was further informed by the genomic analysis of a multidrug-resistant *E. coli* isolate from wastewater in the Virilla River Watershed (Barrantes et al., 2020).

ARG quantifications were assessed using a StepOnePlus™ Real-Time PCR thermocycler (Thermo Fisher, USA). The qPCR standard curve materials were derived from each ARG sequence, and the *16S rRNA* gene was inserted into gBlocks (Integrated DNA Technologies, Coralville, IA, USA), as detailed in Table S3. Calibration curves were generated for each assay, using triplicate replicates of the standard curve material, from 1-log10 to 6-log10 copies/reaction. The relative abundance of each gene was calculated by dividing its absolute abundance by the total abundance of *16S rRNA* genes per sample.

2.5. Shotgun metagenomic sequencing

DNA sample extracts were shipped frozen to Novogene Inc. (Sacramento, CA, USA) for metagenomic library preparation and paired-end 150 sequencing on an Illumina NovaSeq 6000 instrument. On average, over 50 million reads per sample (combining R1 and R2) were obtained, as observed in Figure S1.

Metagenome library preparation was conducted using the ABclonal Rapid Plus DNA Library Kit (ABClonal, Woburn, MA, USA). Each library was evaluated using a Qubit 2.0 fluorometer (Invitrogen, Carlsbad, CA, USA) for preliminary concentration assessment and an Agilent 2100 Bioanalyzer (Agilent, Santa Clara, CA, USA) to determine the insert size.

2.6. Preprocessing and filtering of the raw metagenomic data

The 18 pairs of demultiplexed fastQ files were quality-controlled using FastQC v0.11.9 (github.com/s-andrews). Quality filtered and trimmed them from the Illumina P5/P7 adapters (including poly-G ends) using FastP v0.23.2 (Chen et al., 2018) with non-default parameter -e 25. Human genome contaminants of urban or experimental origins were filtered by aligning the trimmed reads to the UC Santa Cruz hg19 version of *Homo sapiens*, and the resulting reads were rechecked using FastQC v0.11.9.

2.7. GTDB chromosomic and plasmidic genome databases

We conducted a rapid profiling of microbial species diversity using high-quality reads from the Genome Taxonomy Database (GTDB) (Parks et al., 2018), analyzing chromosomal and plasmid genomes separately. This database does not include complete plasmid sequences.

To do this, plasmid sequences from 1742 species (out of 65,703 genomes in GTDB release 207) were isolated from chromosomal sequences. We then created separate indices for these genome collections using Kraken2 v2.0.9 (Wood et al., 2019) and Bracken v2.7.0 (Lu et al., 2017). This allowed us to classify reads independently against the chromosomal and plasmid databases in Kraken2, with species-level counts estimated using Bracken.

For annotation, we used ETE3 (Lex et al., 2014) to gather NCBI taxonomic paths for each species' Taxon ID. This enabled us to annotate plasmid and chromosomal species tables based on read counts per sample. We excluded three mislabeled GTDB-plasmid records from the plasmid dataset, whose lengths were significantly greater than typical chromosomal lengths. These excluded records were *Shinella* sp. PSBB067, *Legionella adelaidensis*, and *Prevotella* sp. oral taxon 299, which appeared in 17, 3, and 2 samples, respectively, with average relative abundances of 0.3 %, 0.02 %, and 0.006 % of reads.

2.8. Diversity analyses of the GTDB chromosomic and GTDB plasmidic profiling datasets

Alpha diversity indices in terms of species richness (number of species), Shannon's entropy (Schmitt and Herzel, 1997), and beta diversity, including the Aitchison distance, were measured in QIIME2 v2023.3 (Bolyen et al., 2019). A rarefaction depth of 3000 reads was chosen to ensure the inclusion of all samples in the diversity analyses. This threshold was determined by the sample with the lowest classified read count (S037, sediment), which yielded 3090 reads after taxonomic classification using the GTDB plasmidic database. While this approach may reduce overall richness estimates, it ensures comparability across all samples and prevents the exclusion of this critical sediment sample. Additionally, rarefaction was applied to control for uneven sequencing depth, a method that demonstrated robustness and reliability even at lower thresholds in simulation-based studies (Schloss, 2024).

Various statistical methods were applied to analyze the data based on the underlying assumptions of each dataset. Normality was assessed, and homogeneity of variances was evaluated before selecting the appropriate statistical method. For datasets that met the assumptions of normality and equal variances (e.g., alpha diversity metrics), one-way ANOVA and Tukey HSD tests were applied to compare groups. For datasets that did not meet these assumptions (e.g., qPCR data, fecal coliforms, *E. coli,* and *E. faecalis*), non-parametric tests such as the Kruskal-Wallis H-test and/or Spearman correlation were used, as they do not require normal distribution. Multiple test corrections were applied where relevant to ensure robust statistical interpretations.

Kruskal-Wallis tests were conducted to evaluate differences in microbiological and physicochemical indicators among the sampling sites. Dunn's post-hoc tests, with Bonferroni correction for multiple comparisons, were applied to identify pairwise differences. The statistical analyses were performed using R language (version 4.3.2.) with the dunn.test package.

Alpha diversity and qPCR-derived ARG abundance correlations were calculated in Python using Spearman's correlation 'scipy.spearman' (Virtanen et al., 2020) and seaborn heatmaps. Statistical tests, including normality tests (Shapiro-Wilk test), Bonferroni-Hochberg corrected *t*-tests, Kruskal-Wallis H, and one-way ANOVA followed by Tukey's Honest Significant Differences (HSD), were realized in Python using 'statannotations,' 'scipy', and 'statsmodels,' as well as metric dimensional scaling using 'sklearn.' Species differential abundances were ranked in term of log-fold change based on a multinomial regression model using a single-layer neural network architecture build by 'songbird' and implemented in QIIME2 v2020.6 (Bolyen et al., 2019). The ranks learning process used an 80:20 train-test split, a batch size of 5, a learning rate $1e^{-4}$, and 100,000 epochs. The model trained to rank species by site and sample type was compared to one for sample type

only and the ratio of errors was reported as a measure of modelling quality (Pseudo Q^2). All plots were made in Python using a combination of 'matplotlib' and 'seaborn' (see code availability).

2.9. Assembly and mapping of all reads and reads classified to GTDB plasmid species

High-quality paired reads were merged for each sample using FLASH v1.2.11 (Magoč and Salzberg, 2011) and co-assembled into contigs for each group of three samples belonging to the same sample type and site using metaSPAdes v3.14 (Nurk et al., 2017). Co-assembly quality was assessed using metaQuast v5.2.0 and the N50.sh script (github.com/hcdenbakker/N50.sh) to measure contig numbers, assembly size, and N50/N90 values. The amount of co-assembled contigs longer than 3 kb that originated from non-prokaryotic genomes was estimated using Tiara v1.0.3 (Karlicki et al., 2022). The reads of each sample used in a co-assembly were mapped to the resulting contigs using Bowtie2 v2.4.4 (Langmead and Salzberg, 2012) and counted using pysam v0.19.1 and SAMtools v1.14 (Danecek et al., 2021; Li et al., 2009). For each sample, mapping rates were calculated concerning the total number of reads in that sample, both for all contig-mapping reads and specifically for those reads classified by Kraken2 to at least one species of the GTDB plasmids database, hereafter referred to as "GTDBp reads." For each contig, the number of reads (and GTDBp reads) mapped to it was divided by the total number of reads in the sample (in millions) and by its length to obtain reads per kilobase per million (RPKM) values for quantitative statistics.

2.10. Detection of plasmidic contigs and circularity

Contig sequences were annotated as plasmidic using manual read tracking and four different software tools: PlasmidFinder with a minimum coverage of 75 % and the KMA mapping method (Carattoli et al., 2014); PlasForest v1.2 with a batch size of 500 (Pradier et al., 2021); Platon v1.7 in "accuracy" and "meta" modes (Schwengers et al., 2020); and PlasX using MMseqs v10.6 (Yu et al., 2024). Manual read tracking involved parsing the Bowtie2 reads-to-contigs mapping files to extract contigs with at least one mapped GTDBp read (see the Profiling section above). These contigs are hereafter referred to as "GTDBp contigs." Contig circularity was identified using "ccfind" (Nishimura et al., 2017) and by detecting "reverse-forward" oriented reads paired over contig ends. This latter approach identified paired reads mapped in reverse-forward orientation, forming a gap longer than the contig length minus three times the median insert size between forward-reverse oriented read pairs (see Method in Yu M.K. et al., 2024). The sets of contigs detected as plasmids and circular by different combinations of these approaches were investigated using the Python implementation of UpSetR (Conway et al., 2017), available at github.com/jnothman/UpSetPlot (Lex et al., 2014).

2.11. Reconstruction and plasmid system networks

Plasmid systems were reconstructed using MobMess (Yu M.K.et al., 2024) for the 673 contigs classified as plasmidic with a PlasX coefficient greater than 0.95, including 42 identified as circular by the "reverse--forward" approach. These systems corresponded to dereplicated clusters of contigs, possibly originating from different co-assemblies and thus from different sample types and sites. The network of plasmid systems was opened in Cytoscape v3.10.2 (Shannon et al., 2003), and the node table was edited to label each node's sample type and site (i.e., set of clustered contigs) located at the ends of directed edges (i.e., plasmid backbone-to-compound containment). These labels, indicating changes in sample type and site, were plotted on the network and used to measure the frequencies of plasmid gene cargo expansions from backbone to compound across the sampling sites and sample types.

2.12. Taxonomy and beta diversity comparisons between GTDBp contigs and profiling

Taxonomy identifiers assigned by Kraken2 to GTDBp reads were compared for each GTDBp contig. Fractions of GTDBp reads unassigned or assigned to one or *Multiple species* were calculated for each sample and their consistency measured from species to phylum levels. The total GTDBp read count was summed for each sample per consistently assigned species to build a feature table and compute an Aitchison distance matrix, which was correlated with that derived from read profiling against the GTDB plasmid database, using the Mantel test (999 permutations). The number, RPKM abundance, length, and circularity of the GTDBp contigs assigned to the 5 % of species found most differentially abundant in the polluted site based on the profiling analyses were investigated across sample types and sites.

2.13. Annotation of antibiotic-resistance genes

Open reading frame (ORF) coordinates and proteins sequences were predicted from each co-assembly contigs using Prodigal v2.6.3 (Hyatt et al., 2010) and annotated to KEGG's KO and Pathways using eggnog-mapper v2.1.9 (Cantalapiedra et al., 2021) based on EggNOG v5.0.2 (Huerta-Cepas et al., 2017) to obtain PFAMs and gene ("Preferred_name") classifications. The predicted genes were also used as input to predict ARGs based on deepARG (Arango-Argoty et al., 2018), whereas the full contig sequences were directly passed to three other ARG-detection software: starAMR v0.9.1v (Bharat et al., 2022) and abricate v1.0.0 (https://github.com/tseemann/abricate), which were both based on ResFinder 2.0 (Bortolaia et al., 2020), and abritAMR v1.0.13 (Sherry et al., 2023), based on AMRFinderPlus v3.10.16 (Feldgarden et al., 2021). The four outputs were combined using hAMRonization v1.1.4 (Datt and Gregorio, 2024) and further harmonized by (average linkage) clustering of the gene coordinates returned by each tool (or by prodigal for deepARG) to obtain a list of ARG classifications for each contig gene across software. We investigated redundancies and discrepancies between the ARG classification terms of the different software to align the annotations of genes in contigs carrying ARGs according to 2, 3, or all 4 software (corrected terms listed in Table S5). The coordinates in the contig sequences for every ARG co-detected by all 4 software and the reads-to-contigs Bowtie2 mapping files were counted using pysam v0.19.1 and SAMtools v1.14 to calculate RPKM values and to build a table of ARGs counts per sample for quantitative statistical analyses (see below).

2.14. Bioinformatics postprocessing, biostatistical analysis and visualization

All metagenomic analysis software ran on the computer clusters SAGA (Sigma2, Norwegian Research Infrastructure Services) and Kabré (CeNAT, CONARE, Costa Rica), where SLURM 23.02.7 (Yoo et al., 2003) jobs were written by 'metagenomix' (github.com/FranckLejzerowicz/metagenomix). Fully documented postprocessing scripts and Python jupyter notebooks, as well as configuration files (databases, pipeline, modules, non-default parameters) to generate the commands for all analysis software and replicate analyses are openly available at gitlab.com/alper1976/marmip/-/tree/main/keilor/papers/virilla_river/metagenomes.

3. Results

3.1. ARGs are enriched along the pollution gradient and especially in the water column

At three sampling sites along the Virilla River (Fig. 1A), biological and chemical measurements taken during the dry season of 2022 indicate an evident degradation in surface water quality. Upstream Site 1

shows fair quality, while downstream Sites 2 and 3 exhibit poor quality (Fig. 1B, Table S1). A Kruskal-Wallis test indicated significant differences in microbiological and physicochemical indicators among the sampling sites (for fecal coliforms H test: statistic = 7.3, $p = 0.025$; for *E. coli* H test: statistic = 6.9, $p = 0.030$; for *E. faecalis,* H test: statistic = 7.3, $p = 0.026$ and for oxygen saturation, H test: statistic = 8.0, $p = 0.018$). Dunn's post-hoc analysis revealed significant differences in fecal contamination between Site 1 and Site 3 (adjust p-value=0.007), supporting this decline in surface water quality.

This observed decline in water quality was mirrored by the patterns of ARG contamination, underscoring a relationship between anthropogenic pollution and the microbial and genetic composition of the river.

This degradation was accompanied by contamination with ARGs, whose abundances were quantified in water and sediment samples using standard qPCR for 10 genes, normalized to the *16S rRNA* gene abundance (Fig. 1C). The abundances of all ARGs varied significantly among sites (Kruskal-Wallis H test: statistic = 12.1, $p = 0.002$). They were influenced by sample type for individual ARGs, with higher levels generally observed in water samples. The most abundant ARGs were *sul-2, sul-1, tet(A),* and *tet(Q),* with the latter three showing notably higher concentrations in water samples compared to sediment at Site 2 (Fig. 1C).

Interestingly, the less abundant genes bla_{TEM}, *ermB*, and *qnrS* were also enriched in water samples at the polluted Site 2. A similar enrichment was observed for *intI-1* at the most contaminated site (Site 3). The *intI-1* gene displayed a clear enrichment in sediment samples along the pollution gradient, as confirmed by the Tukey HSD test (Site 1 vs. Site 2, $p = 0.0176$; Site 1 vs. Site 3, $p = 0.0016$). This trend was also evident for the two sulfonamide resistance genes (*sul-1* and *sul-2*) in comparisons between Site 1 and Site 3 across both sample types, with *sul-1* exhibiting slightly higher mean differences (Tukey HSD: mean difference = 0.075, $p < 0.003$) than *sul-2* (Tukey HSD: mean difference = 0.02, $p = 0.015$).

Three ARGs were undetected, but only in the least polluted sediment samples: bla_{TEM} (at sites 1 and 2), *dfrA12*, and *qnrS* (at site 1). Additionally, $bla_{CTX\text{-}M}$ was not detected in any of the samples. In summary, ARGs were enriched along the gradient of anthropogenic activities, with water samples showing the highest contamination levels.

The varying enrichment patterns of specific ARGs, such as bla_{TEM}, *ermB*, and *qnrS*, suggest a nuanced interaction between pollution sources and the ecological dynamics of the river's microbial communities. These findings set the stage for examining diversity trends in plasmid and chromosomal species.

3.2. Plasmidic species diversity increases with biological and chemical pollution and differences in chromosomic diversity depend on sample type

Alpha and beta diversity analyses of species profiles revealed

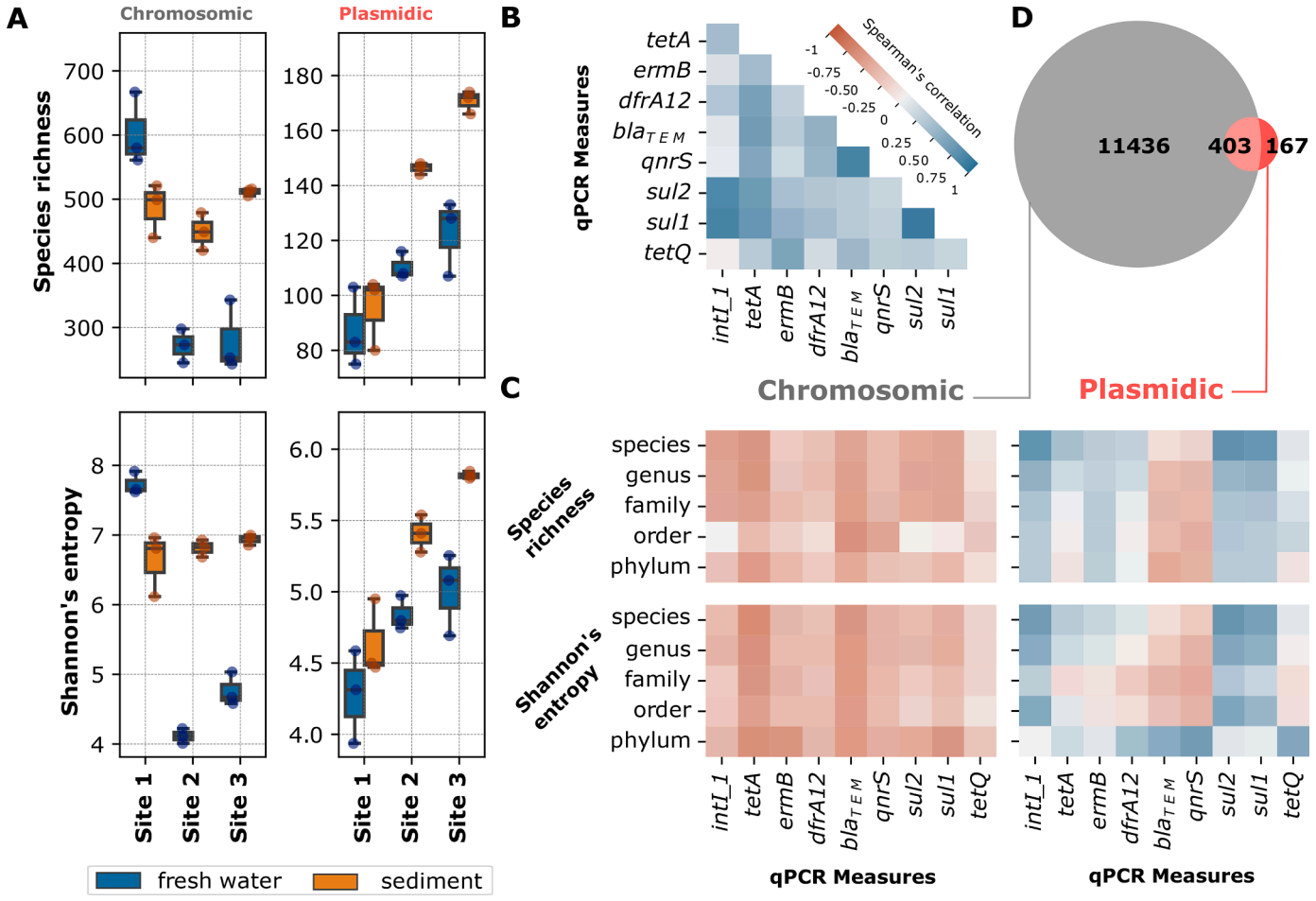


Fig. 2. Species richness and Shannon's entropy alpha diversity of the GTDB chromosomic and plasmidic profiles. (A) Indices values for each site and sample type (significant relationships between these variables are detailes in the main text) (B) Spearman's correlations amongst qPCR-derived ARG abundances (upper triangular matrix) and (C) between these ARG abundances and alpha diversity indices computed for the chromosomic and plasmidic datasets collapsed at different taxonomic levels (lower matrices). Indices were measured of all three samples per dataset, site, and sample type after rarefaction to 3000 reads per sample. (D) Number of classified species in the chromosomic (grey) and plasmidic (red) profiles. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

contrasting patterns in response to pollution and sample type for metagenomic reads classified against the GTDB chromosomic or plasmidic databases. Indices were measured of all three samples per dataset, site, and sample type after rarefaction to 3000 reads per sample.

These diversity analyses highlight that pollution gradients increase plasmidic diversity and create distinct ecological niches, shaping microbial community structure in water and sediment samples differently.

Plasmidic alpha diversity increased with anthropogenic pollution (Fig. 2A), reflected in both species richness (570 species, Kruskal-Wallis H-test: $H = 11.87$, $p = 0.003$) and Shannon's entropy ($H = 10.05$, $p = 0.006$). This increase was more pronounced in sediments, where richness showed a tenfold higher F statistic than water (one-way ANOVA: $F = 66.85$, $p = 0.0001$, versus $F = 6.99$, $p = 0.027$).

In contrast, the richness of chromosomic species (11,839 species) halved in water samples, decreasing from an average of 602.7 (± 56.5) at upstream Site 1 to a more stable 275.8 (± 38.9) at polluted Sites 2 and 3 (Tukey HSD tests vs. Site 1, both $p = 0.0004$, Fig. 2A). Shannon's entropy for chromosomic data remained consistent across sediment sites (one-way ANOVA: $F = 0.96$, $p = 0.43$).

Overall, alpha diversity at polluted sites was consistently higher in sediments than in water, with *t*-test statistics averaging 7.5 (± 0.35) for chromosomic richness and 8.7 (± 4.16) for plasmidic richness (all Benjamini-Hochberg adjusted p-values < 0.01).

Note that these patterns are not due to differences in read amounts (Figure S1A), which do not show differences across site or sample type (Figure S1B and S1C). Fractions of classified reads were lower at Site 1,

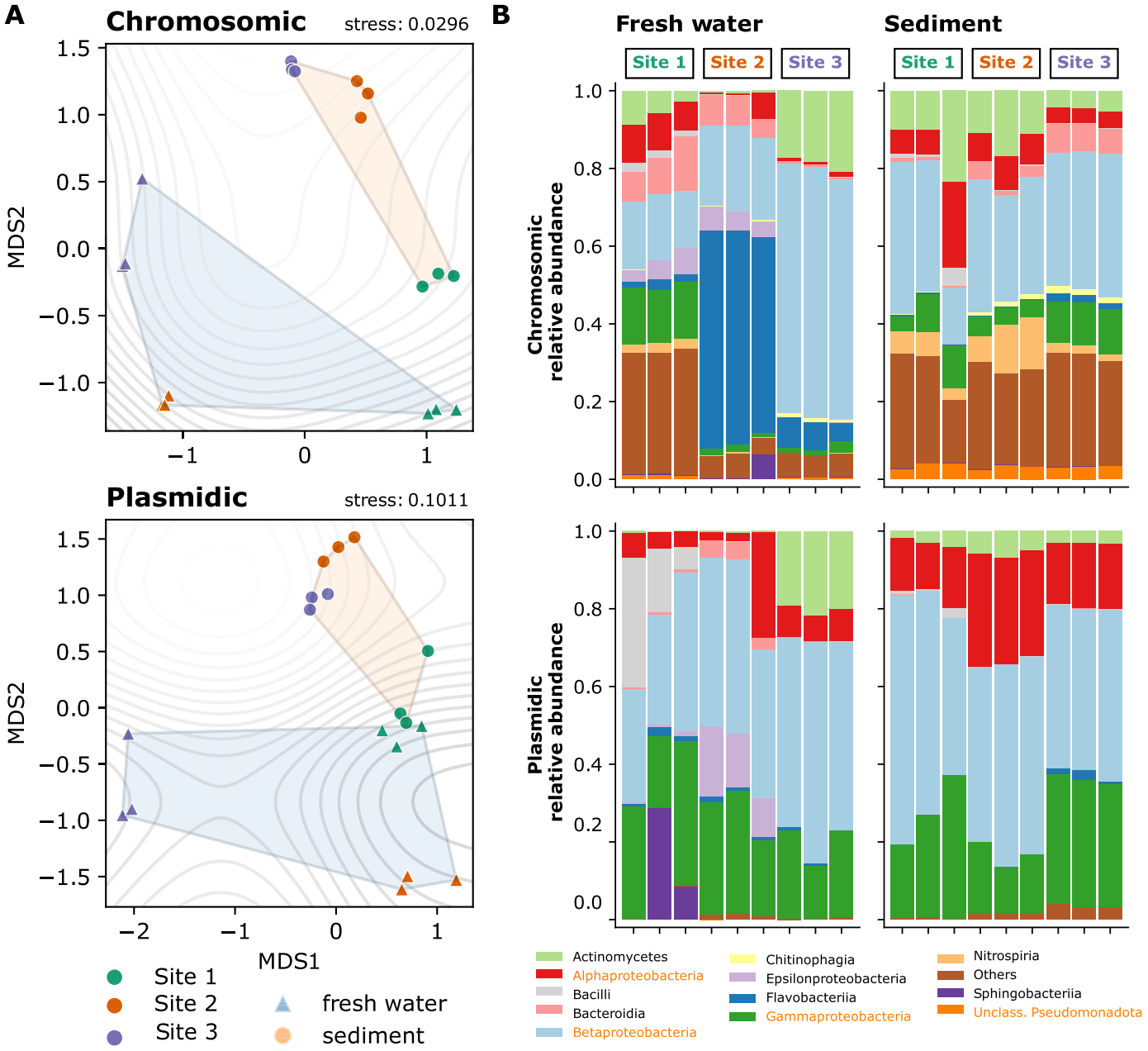


Fig. 3. Beta diversity structure and taxonomic composition of GTDB species profiles. (A) Metric multidimensional scaling ordination of the chromosomic (upper panel) and plasmidic (lower panel) Aitchison distance matrices, with site-colored samples delineated by convex hulls for water (blue area triangles) and sediment (orange area dots) sample type. Isolines represent a smoothed surface fitted by a generalized additive model for site. (B) Relative abundances of reads classified to the chromosomic (upper panels) and plasmidic (lower panels) genome sequences of the GTDB databases, colored at class level (Pseudomonadota phyla written in orange) for each of the three water (left panels) and sediment (right panels) samples per site. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

where the chromosomic dataset was most diverse (Figure S1C). We observed that the diversity of plasmid species increased with pollution, contrary to chromosomic diversity, which exhibited different trends depending on the sample type. The *intI-1* gene showed enrichment in sediment samples along the pollution gradient, supporting the hypothesis that urbanization-induced pollution impacts microbial community composition. The GTDB chromosomal profiles showed contrasting trends to the plasmidome, with reduced chromosomal diversity observed in water samples at polluted sites, likely driven by ecological filtering favoring specific taxa.

ARG abundances were generally positively correlated with each other. The gene *intI-1* exhibited a strong positive correlation with *sul-1* and *sul-2*, commonly associated with mobile genetic elements and resistance to sulfonamide antibiotics, respectively. The genes *intI-1, sul-1*, and *sul-2* demonstrated a significant negative correlation with chromosomal alpha diversity metrics (Shannon's entropy and species richness), indicating reduced chromosomal diversity in environments enriched with ARGs. Interestingly, *sul-1* and *sul-2* displayed weaker correlations with less abundant ARGs such as bla_{TEM}, *ermB, qnrS*, and *dfrA12*. These findings underscore the ecological filtering effects of urbanization-induced pollution, the pivotal role of *intI-1* in ARG dissemination, and the associated reductions in chromosomal diversity along pollution gradients (Fig. 2B). This pattern was consistent across indices computed at every taxonomic level. Notably, *intI-1* and the sulfonamide resistance genes positively correlated with plasmid diversity, which increased at the species level with pollution in both sample types. Interestingly, 167 (29.3 %) of the 570 plasmid species (representing 23 genera) could not be classified using GTDB chromosomes (Fig. 2C). These unclassified species accounted for an average of 9.7 % (± 8.3 %) of metagenomic reads across the 18 samples, with a higher proportion in Site 1 (23.6 ± 12.4 %) compared to the more polluted sites (3.8 ± 2.3 %).

3.3. Sample type also influences plasmidic profiles, but the abundance of potential plasmid hosts changed consistently in water and sediments

Key taxa composing chromosomic and less diverse plasmidic profiles might explain differences in beta diversity (Fig. 3). Metric multidimensional scaling (MDS) of the compositionally robust Aitchison distance matrix revealed that plasmidic and chromosomic species communities clustered by sample type and site (Fig. 3A).

Plasmidic beta diversity among Site 1 samples appeared less variable than for polluted Sites 2 and 3, in which sediment and water samples clustered separately (PERMANOVA respectively explained 36, 18, and 16 % of variance by site, sample type, and their interaction, all p-values < 0.003). There were differences in dispersion, except within sample types for the chromosomic profiles (PERMDISP, F-value=0.6, p-value=0.286, 999 permutations). In the sediment, mean distances among samples of Site 1 in both datasets were notably higher than among samples of Site 3. Still, for the plasmid dataset, their variation tended to decrease along the pollution gradient, while it tended to increase for water samples (Figure S2).

These findings further emphasize the role of pollution gradients in shaping plasmidic and chromosomic community structures, with distinct clustering patterns revealing the combined influence of sample type and site-specific factors.

These subtle patterns of chromosomic and plasmidic beta diversity differences reflect major differences in taxonomic composition. Overall, Pseudomonadota was the most abundant phylum, ranging from 35 to 85 % in water and 42 to 94 % in sediment samples, and notably class Betaproteobacteria in polluted sites (Fig. 3B). Interestingly, class Flavobacteriia exclusively dominated in the water samples of Site 2, for the chromosomic dataset (67 to 71 %). For the plasmid dataset, it was possible to rank microbes differentially abundant in the polluted site using Site 1 as a reference in a multinomial regression model also accounting for sample type (Pseudo Q^2=0.413, Figure S3). Thirteen species enriched in the most polluted Site 3 belonged to Actinomycetes (all Mycobacteriales). Strikingly, in both sample types, the abundance of these different Site 3-enriched species increased from Site 1 to Site 3, even when grouped at class level (Figure S3B), suggesting a plasmid host spectrum limited to these classes. However, the abundance of species enriched in Site 2 did not necessarily increase from Site 2 to Site 3, which for 2 Betaproteobacteria and the 9 Gammaproteobacteria species tended to stabilize (Figure S3C).

3.4. Reads classified to GTDB plasmids assemble into long contigs that prevail in polluted water samples and relate to pollution-enriched species

Metagenomic reads classified as GTDB plasmids assembled into long contigs that were more prevalent in polluted water samples and associated with species enriched at these sites. Water sample contigs were generally longer than those from sediments, and their N50 values increased with pollution levels (Tables S5–S8). GTDB plasmid contigs represented a higher fraction of co-assemblies at polluted sites (30 % at Sites 2 and 3) compared to the upstream Site 1 (5.4 %), regardless of sample type (Figure S4A). High-abundance contigs (≥2 kb with >90 % read mapping) were larger and most frequent at Site 3, particularly in water samples, and were associated with taxa such as *Mycobacterium* sp., *YC-RL4, Methylomonas*, and *Enterobacter* (Figure S5).

Correlation between contig-based and read-based diversity profiles (Mantel test, Spearman rho = 0.756, $p = 0.001$) further validated these patterns. Although plasmid reads accounted for only 0 .15 % of all mapped reads, they represented 3 4.5 % of those contributing to co-assemblies (Table S10). These results emphasize the role of pollution in driving the assembly of long plasmidic contigs, highlighting their association with resistant taxa and the intensifying impact of anthropogenic pressures on microbial communities along the pollution gradient.

3.5. Circular contigs also relate to pollution-enriched species, and consistently annotated plasmids originate from polluted sites

Of all co-assembled contigs, 42 were identified as circular based on reverse-forward oriented reads, with 27 also detected by ccfind (Fig. 4A). The identification of circular contigs provides additional evidence for the mobilization and persistence of plasmidic elements in polluted environments, offering further insights into the genetic mechanisms at play. The ccfind method identified an additional 856 circular contigs, consistently more than the reverse-forward method across co-assemblies (Fig. 4A). The RPKM abundance of the 42 circular contigs was higher in freshwater and polluted samples, similar to linear contigs but over a smaller range. In contrast, the 856 ccfind contigs were more abundant in sediment and Site 1 samples (Figure S6B), likely reflecting size differences, as 85.7 % of reverse-forward contigs were >3 kb compared to only 29.8 % of ccfind contigs (Table S11). Larger ccfind contigs were observed in sediment from Site 1, with third-quartile sizes of 2774.7 bp at Site 1 and 35,148.2 bp at other co-assemblies.

Five GTDBp circular contigs contained reads classified to six of the 28 species enriched at polluted Site 3. The longest of these (24.5 kb) mapped reads from two *Mycobacterium* species, with related GTDBp reads also mapping to eight linear contigs classified as *Mycobacterium grossiae*. This circular contig included a nitrogen reductase and a NO3/NO2- transporter operon flanked by conjugation (TrwC relaxase) and transposition elements (Fig. 4C). Two other circular contigs from uncharacterized *Mycobacterium* species enriched at Site 3 were smaller (1.3 and 1.7 kb) and contained minimal genes (Figure S7). GTDBp reads classified to these species also mapped to 112 linear contigs, potentially part of these circular genomes.

Plasmid detection used four methods (Plasmidfinder, PlasX, Platon, and PlasForest; Figure S8). The largest plasmid set (1151 contigs) was unique to PlasForest and GTDBp, with these methods identifying short plasmids, predominantly from polluted sites, with little evidence of

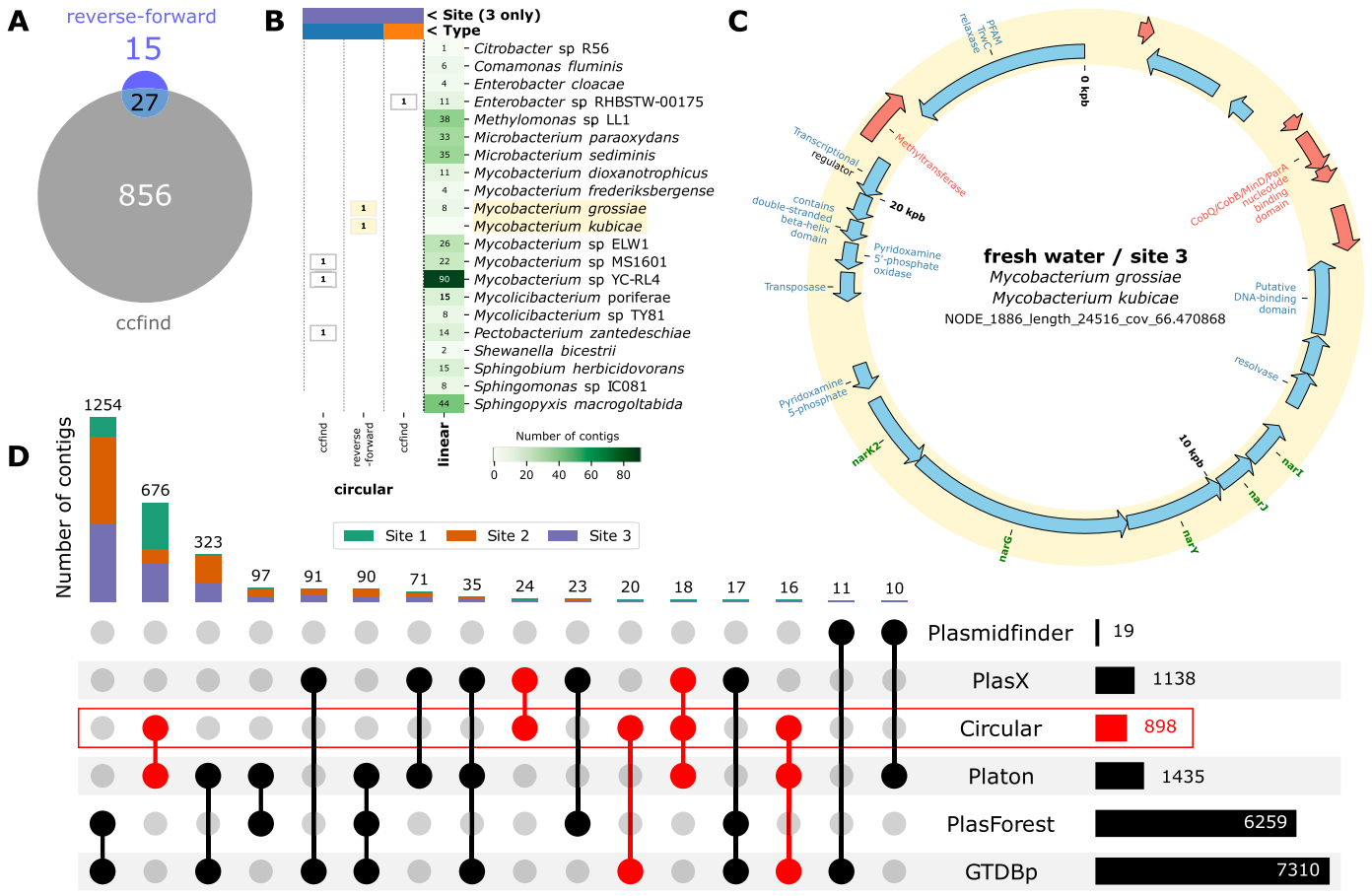


Fig. 4. Annotation and circularity of plasmidic contigs and GTDBp contigs. (A) Number of plasmids detected as circular by "ccfind", the "reverse-forward" approach, or both (see Yu et al. 2024). (B) Number of circular (boxed) and linear (green) GTDBp contigs made of GTDBp reads classified to the 5 % of species found most differentially abundant in site 3 (see profiling results). One circular contig is made of GTDBp reads classified to both Mycobacterium grossiae and M. kubicea (highlighted in yellow). (C) Gene map of the circular GTDBp contig made of GTDBp reads classified to M. grossiae and M. kubicea. Genes in red and blue are on forward and reverse strands, respectively, and the green names indicate nitrogen metabolism genes for a nitrate reductase (narI: gamma subunit; narJ: delta subunit; narY: beta subunit; narG: alpha chain) and a major facilitator superfamily (narK2: probable nitrate or nitrite transporter). (D) Sets of at least 10 contigs co-detected by multiple plasmid-detection tools, incl. the GTDBp read tracking approach and detection as circular (see Figure S11 for all sets). Contigs intersection sets are indicated by connected dots (lower panel) and their sizes by bars colored per site (upper panel, with number of contigs indicated). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

circularity (Figs. 4D and S9). These results illustrate the diversity of plasmidic elements in polluted environments, highlighting their potential roles in facilitating resistance gene transfer and adaptation. Another large set (676 contigs) included 47.1 % of Platon-predicted plasmids and 75.3 % of circular contigs, with 98.2 % confirmed as circular by Platon. Of the 42 circular contigs identified by reverse-forward reads, only 20 overlapped with this set.

3.6. *Plasmids acquire resistome genes along the gradient and transposable metabolic traits associated with the pollution-enriched genus*

The 673 contigs annotated as plasmids with a score >0.95 by PlasX formed 31 networks (Figure S10A), including 77 compound contig nodes connected by 47 directional edges representing gene cargo transfers from shorter to longer contigs. These findings reveal the dynamic nature of plasmid evolution in polluted sites, with compound plasmids playing a key role in transferring resistance and metabolic traits. Three networks contained compound plasmids derived from two distinct backbones, all observed at Site 3. Notably, 44.6 % and 69 % of the cargo genes occurred within the same site or sample type, respectively, with 92.8 % of plasmids involved in compound formation co-assembled from polluted Site 2 and Site 3 water samples (Figure S10B). Compound plasmids were more abundant in water samples than their backbones at upstream Site 1, while in sediment, the log-ratio of compound plasmids to backbones increased with pollution. This trend was marginally significant between sites 2 and 3 (Kruskal-Wallis test, $H = 3.85$, adjusted p-value=0.049) (Figure S10C).

One of the two largest plasmid networks demonstrated extensive gene cargo enrichment between Site 2 and Site 3, culminating in forming a large compound plasmid (23.8 kb) at Site 3. This plasmid included genes for replication, conjugation, type II and IV secretion systems, transposases, aminoglycoside resistance, and an operon for mercury scavenging, transport, and regulation (Figure S10D, upper part). Increased plasmid diversity and complexity were observed alongside the rise in ARGs. The second large network consisted solely of contigs from water samples at Site 3 and included three compound plasmids. The longest of these was circular and encoded a transposase, a conjugative relaxase, two resolvases, and 10 other genes, with all 10 belonging to *Mycobacteriaceae* orthologs out of a total of 19 genes (Figure S10D, lower part).

3.7. *ARG enrichment along the pollution gradient largely conveys aminoglycoside and sulfonamide resistance*

On average, 1.56 (± 0.49) million open reading frames (ORFs) were predicted per co-assembly (Table S11). Still, only 1513 ARGs were

detected, largely (78.1 %) by deepARG, which operates on ORFs, while abricate, abritAMR, and starARM operate on contigs (Figure S11A and Table S13). The significant enrichment of ARGs in polluted sites highlights how anthropogenic pressures, particularly wastewater discharge, drive the proliferation of resistance mechanisms in microbial communities. The number of ARGs detected by at least two annotation tools was reduced by an order of magnitude, highlighting the conservative nature of multi-method validation (Figure S11B). Those ARGs belonged to aminoglycosides, beta-lactams, MLS, or sulfonamide (Figure S11C). Irrespective of the stringency regarding method consistency, the number of co-detected ARGs was systematically higher in contigs co-assembled in highly polluted sites but similar across types (Figure S11D). Annotations were mainly consistent for the major ARG classes (aminoglycosides, beta-lactams, sulfonamide, tetracycline) that gathered 71.4 % of the predicted genes. Yet, for 17.5 % of the ARGs, deepARG only employed the term MLS, which was used to group annotations by other methods to macrolide, lincosamide, and/or streptogramin b, but never the term diaminopyridine, which was grouped as folate pathway antagonist as per majority rule (Figure S11E). These results underline the complexity of ARG classification, emphasizing the need for multiple

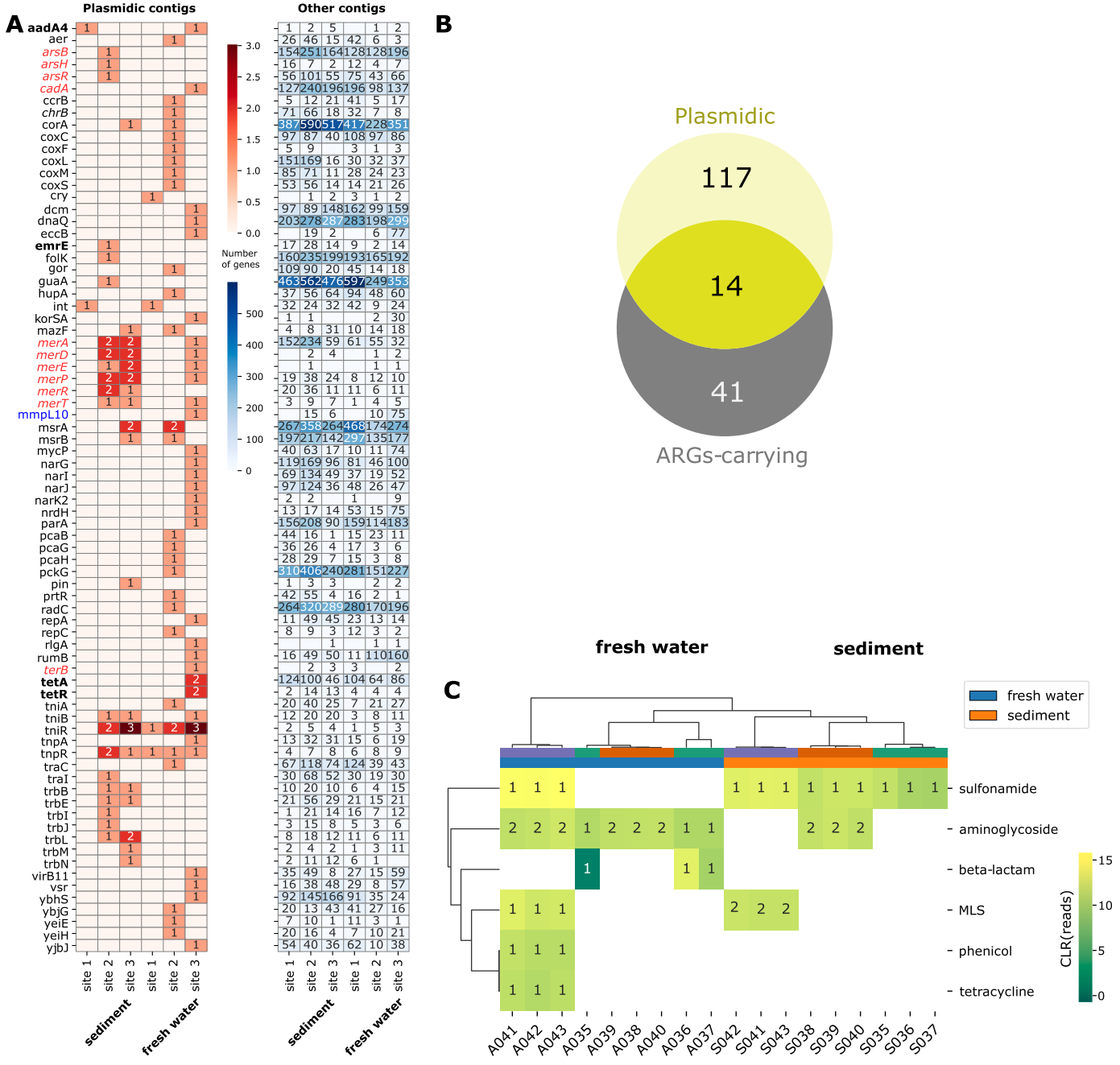


Fig. 5. Number and abundances of plasmid-carried genes and ARGs. (A) A number of genes predicted on plasmidic (left panel) and non-plasmidic (right panel) contigs for each co-assembly, for each of the gene names annotated by eggnog-mapped ("Preferred_name" classification) and present on plasmidic contigs (more genes have been classified in non-plasmidic contigs). Names in bold, red, and blue code for proteins associated with antibiotic resistance, metal metabolism, and *Mycobacterium*, respectively (B) Total number of contigs that are plasmidic (as per ≥ 3 approaches), that carry ARG(s) (as per all 4 software), or both. (C) Clustered heatmap of the CLR-transformed plasmid-carried ARG read counts (numbers of genes annotated) for each final ARG class (rows) and samples (columns). For all plots, plasmidic contigs are those co-detected by at least 3 (out of the 5) different plasmid-detection approaches, and ARGs are those co-detected by all 4 ARG-detection software. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

annotation approaches to capture the full spectrum of resistance traits present along the pollution gradient.

Of the 63 ARGs co-detected by all four methods, only five were found at the upstream Site 1 (Figure S11F). These included three aminoglycoside ARGs in water samples, which were significantly less abundant than those at Site 3 (Kruskal-Wallis test, $H = 11.3$, adjusted p-value=0.008), and a single sulfonamide resistance gene.

This disparity between upstream and downstream sites underscores the cumulative impact of anthropogenic activities, with Site 3 emerging as a critical hotspot for ARG accumulation and dissemination.

The sulfonamide resistance gene was present in two copies at each polluted site co-assembly but reached its highest abundance at Site 3 (Kruskal-Wallis test vs. Site 2, $H = 10.4$, adjusted p-value=0.013) (Figure S11G), particularly in ARG-rich water samples (Figure S11F).

Similarly, GTDBp contigs annotated to five out of seven ARG classes were exclusively found in water samples from polluted sites. Among these, 14 contigs carried MLS, sulfonamide, or tetracycline resistance genes, and one contig each for folate pathway antagonist and phenicol resistance, all observed only at Site 3. Their RPKM abundances were systematically higher in this highly polluted site (Figure S12). These results underscore the selective pressures exerted by pollution and emphasize the role of water as a medium that facilitates ARG enrichment and spread.

Increased abundances of ARGs at polluted sites were also observed in sediment, where only one GTDBp contig carrying a sulfonamide resistance gene was detected at Site 1.

3.8. Plasmidic genes related to aminoglycoside, sulfonamide, and metal resistance abound in water and polluted samples

Among the 77 genes identified through holistic annotation with the EggNOG5 database, 4 were related to antibiotic resistance (*aadA4, emrE, tet(A)*, and *tet(R)*) and 12 to metal resistance (arsenic, cadmium, mercury, and tellurium), all exclusively in polluted sites (Fig. 5A).

Notably, six mercury resistance genes were primarily found on plasmids in sediment samples, although non-plasmid copies also occurred at polluted sites. A plasmid copy of the *mmpl10* gene, coding for the *Mycobacterium* membrane protein large 10, was detected in the sediment of Site 3, while non-plasmid copies were found only at contaminated sites (Fig. 5A).

Of the 131 contigs identified by at least three plasmid-detection methods, 14 were consistently predicted to carry at least one ARG by all four annotation methods. Together, these contigs encoded 17 ARGs (Fig. 5B).

CLR-transformed read counts showed the overall distribution of plasmidic ARG abundances. Plasmids from water samples at Site 3 harbored a diverse range of ARG classes. In contrast, those from sediment samples carried a more limited set, including sulfonamide and MLS resistance genes at Site 3 and aminoglycoside resistance genes at Site 2 (Fig. 5C).

Interestingly, 8 ARGs were carried on plasmids corresponding to GTDBp contigs, and these mapped reads were classified to three of the 28 species found differentially abundant at Site 3. One plasmid related to *Enterobacter* sp. *RHBSTW-00175* carried two MLS copies in Site 3 sediment, while three other plasmids from polluted sites were associated with *Pectobacterium zantedeschiae* and carried three beta-lactam and one aminoglycoside resistance genes. Circular GTDBp contigs were also associated with these species (Figures S7 and S13). For *Shewanella bicestrii*, two GTDBp contigs predicted as plasmidic were identified: a short contig (757 bp) carrying a beta-lactam resistance gene and a longer one (6 kbp) carrying both an aminoglycoside and a sulfonamide resistance gene, both in water samples from Site 3 (Table S14).

These associations between plasmid-encoded ARGs and pollution-enriched species underscore the complex interactions between microbial taxa and plasmid-mediated resistance mechanisms in polluted environments. This suggests that plasmid genes conferring resistance to aminoglycosides, sulfonamides, and metals are highly prevalent in polluted water samples, potentially equipping their hosts with multi-resistance capabilities.

4. Discussion

4.1. Impact of pollution on microbial diversity and community structure

We demonstrate the differential impacts of pollution on plasmid and chromosomal diversity along a tropical riverine system in Costa Rica, despite using only triplicates per treatment in our full-factorial (habitat type and pollution level) gradient study. As urban pollution levels intensified, plasmidome and ARG diversity tended to rise, while chromosomal diversity decreased in the water column or remained unchanged in the upper sediment. These findings align with previous studies on antimicrobial residues and ARGs detected in the surface water of the Virilla River Watershed, including analysis of the plasmids from bacterial isolates (Mendoza-Guido et al., 2024; Morales-Mora et al., 2025; Ramírez-Morales et al., 2021; Wilkinson et al., 2022).

Anthropogenic pressures such as the release of antibiotics and other pollutants from sources such as wastewater treatment plants (WWTPs), landfills, and agricultural runoff, not only enhance the selection pressure on microbial communities, fostering the spread of ARGs but also promote the persistence and mobility of these resistance determinants across different environmental compartments, thereby posing a substantial risk to public health (Czatzkowska et al., 2022). This phenomenon is further exacerbated by the ability of MGEs to cluster ARGs into resistance islands within conjugative plasmids that dominate in multidrug-resistant bacterial populations (Vale et al., 2022; Wang and Dagan, 2024).

In this context, heavy metals and pharmaceuticals contribute significantly to selective pressures, driving increased plasmid diversity. For example, mercury exposure has been shown to elevate the abundance of plasmids carrying mercury resistance genes, similar to the result obtained in our study, highlighting plasmids' critical role in enabling microbial communities to adapt to mercury stress (de Lipthay et al., 2008). Furthermore, HGT is a pivotal mechanism for disseminating adaptive traits among bacteria, especially in polluted environments. As mobile genetic elements, plasmids facilitate the transfer of genes that confer resistance to pollutants, enhancing bacterial adaptation and survival. Studies have shown that plasmids can carry genes that enable bacteria to cope with toxins and pollutants, with gene transfer occurring across phylogenetically and geographically distant bacteria (Petersen et al., 2019). Emerging pollutants, such as microplastics, have also been found to promote plasmid transfer. Microplastics provide surfaces for biofilm formation, increasing plasmid transfer frequency among bacterial populations (Arias-Andres et al., 2018). Environmental stressors, including heavy metals and nutrient pollution, can influence plasmid-mediated gene transfer. For instance, when combined with carbon and nitrogen enrichment, arsenic stress has been shown to enhance HGT in coastal waters (Neethu et al., 2022).

Also, the decreased chromosome diversity at the highly polluted sites can be associated with the adverse effects of pollution on most bacterial species and the proliferation of a few, such as copiotrophs, that gain competitive advantages under such conditions. The increased abundance of these species and the decline of more sensitive, oligotrophic bacteria suggests that environmental filtering under severe pollution reduces ecosystem stability and resilience, affecting microbial networking complexity (Larsson and Flach, 2022; Zhang et al., 2022). Although previous studies have extensively addressed the effects of pollution levels on microbial diversity in aquatic environments (Numberger et al., 2022), we show that an urbanization-induced pollution gradient affects the community plasmidome and its gene content in a riverine ecosystem. Increasing the genetic diversity of the community plasmidome reveals a striking contrast to its chromosomal counterpart.

4.2. Taxonomic resolution of plasmids in polluted environments

Despite their inherent mobility across taxa, the taxonomy assignment to plasmid contigs offers profound insights into their ecological roles, potential hosts, and evolutionary history. This taxonomic resolution is essential for understanding the niches plasmids occupy and their contributions to microbial community functions across varied environments. Determining the likely host range of plasmids allows researchers to infer ecological interactions and potential HGT pathways, which is critical for comprehending the dissemination of antibiotic resistance and other adaptive traits.

Our findings reveal that taxonomic resolution allowed us to characterize shifts in plasmid diversity and composition across the pollution gradient, providing insights into how environmental conditions influence the structure of the plasmidome. Other research has shown the significance of plasmids in microbial adaptation, especially in extreme settings like deep-sea hydrothermal vents, where they influence microbial community structures and enable unique adaptive strategies (Ciuchcinski K. et al., 2024). Incorporating taxonomic data into plasmid analysis deepens our understanding of ecological and evolutionary processes while informing strategies for managing microbial dynamics in diverse ecosystems.

Assigning taxonomy to plasmid contigs advances ecological and evolutionary studies by offering a more detailed understanding of plasmid-mediated gene transfer. This approach is particularly significant in polluted environments, where plasmids act as crucial vectors for resistance gene dissemination and other adaptive traits, emphasizing the importance of ongoing research into their ecological and evolutionary dynamics.

4.3. Plasmid dynamics and ARG proliferation in polluted environments

The consistency in plasmid diversity patterns observed across four independent annotation tools (PlasmidFinder, PlasForest, PlasX, and Platon) underscores the robustness and reliability of our metagenomic workflow for plasmid detection and characterization. Previous research has demonstrated that standardized metagenomic approaches yield reproducible microbial diversity profiles across spatial and temporal gradients (Wallace et al., 2018).

Notably, our results suggest that polluted urban environments are associated with the assembly of longer DNA contigs, potentially due to the selection of plasmids carrying multiple gene clusters, including those encoding ARGs and conjugation/transposition machinery (Alonso-del Valle et al., 2021; Di Cesare et al., 2022). Furthermore, the persistent presence of these longer contigs assigned to genera such as *Mycobacterium*, particularly at polluted sites, aligns with previous observations indicating that specific bacterial taxa become more prevalent in such environments, primarily due to their robust resistance mechanisms. This is corroborated by detecting a circular contig containing the *NarGHJI* operon, which encodes the complete reduction of nitrate to ammonium via nitrate reductase. These genes have been described in other *Mycobacteria* species, including saprophytic soil genera such as *M. smegmatis* and significant pathogens like *M. tuberculosis, M. bovis*, and *M. avium* (Amon, Titgemeyer, and Burkovski, 2009).

In environmental *Mycobacteria*, the presence of these genes on a circular contig flanked by conjugation and transposition elements suggests the potential horizontal transfer of this operon to other bacterial genera, enabling the utilization of nitrate as a nutrient in environments heavily contaminated with these compounds.

Nitrate reduction is a critical process in the nitrogen cycle with significant agricultural, environmental, and public health implications. However, nitrate is also a major pollutant in river waters, with high levels of nitrates and nitrites previously reported in the surface waters of the Virilla River Watershed. This contamination is likely driven by industrial and domestic wastewater discharges and chemical fertilizers used in agricultural systems (Herrera-Murillo et al., 2019; Pérez-Gómez et al., 2021).

However, detecting genes, such as the mycobacterial membrane protein *mmpl10* in non-plasmid copies, indicates that both plasmid-borne and chromosomal adaptations could be instrumental for microbial persistence in contaminated environments. This example underscores the dual mechanisms of genetic adaptation bacteria might employ to survive under pollution-induced environmental stress (Larsson and Flach, 2022).

The higher percentages of compound plasmids, consisting of a backbone and additional cargo genes, suggest plasmids' capacity to acquire or lose genes, particularly in polluted environments. This underscores the potential for mixed anthropogenic factors to accelerate plasmid evolution (Yu M.K.et al., 2024). Moreover, the observed presence of complex plasmid networks underscores the high adaptability and the coevolution of specific genera to carry them more effectively, particularly in response to environmental pressures such as pollution (Stalder et al., 2020). This emphasizes that the environmental release of fecal bacteria may boost the evolutionary process by providing genetic elements adapted to capture and transfer ARGs (Larsson and Flach, 2022). Remarkably, many contigs assembled in our study remained unclassified at the species level, highlighting the limitations of metagenomic approaches, the constraints of current databases like GTDB, and the potential for discovering novel microbial taxa or plasmids in polluted environments.

This divergence in plasmid and chromosomal diversity trends was detectable because plasmid sequences from each GTDB species were extracted to build a dedicated profiling database, highlighting a potential limitation for studies that profile both genome types indiscriminately. However, given that the GTDB plasmid database includes only 1742 species, it likely underrepresents the true plasmid diversity and lacks records of plasmids that can be shared across distantly related taxa. This limitation may also explain why a third of the species detected in plasmid profiles were absent from chromosomal profiles, potentially influencing the observed plasmid diversity patterns.

We re-ran the read profiling against the GTDB plasmid database to validate this pattern. However, this time using only the subset of reads that mapped onto contigs confidently identified as plasmidic by PlasX (score >0.9). The resulting diversity trend remained consistent with that shown in Figure S14, further supporting the robustness of our approach.

Moreover, in these highly plasmid-enriched polluted sites, multiple Mycobacteriales species were consistently abundant, suggesting plasmids proliferate within this taxonomic group, potentially facilitated by HGT. This interpretation is further supported by the reconstruction of plasmids carrying genes annotated to Mycobacteriales, onto which reads classified as Mycobacteriales were also mapped.

Additionally, in previous work using sediment samples from the Virilla River, we benchmarked various plasmid identification tools. While each tool has distinct strengths, no single method is comprehensive (Rojas-Villalta et al., 2023). This benchmarking reinforces our multi-tool approach in this study and highlights the need for combining methods to achieve robust plasmidome characterizations in complex environmental samples.

While we acknowledge the limitations of relying solely on metagenomic data, the diversity of identified plasmids and their associated functional cargo provide valuable ecological insights into plasmid dynamics across the pollution gradient. These limitations underscore the need for further research and methodological improvements, including database expansion. A critical avenue is understanding the mechanisms driving plasmids' diversity patterns and chromosomal vs. plasmids diversification, including antimicrobial resistance spread under anthropogenic pressures. The former includes how plasmids' dispersal, drift, and selection interact to promote and suppress antimicrobial resistance spread.

4.4. Environmental factors driving ARG dissemination in water and sediment

The significant variation in ARG abundances among sites underscores the spatial heterogeneity of ARG distribution in environmental matrices. The observation that water samples exhibited higher ARG levels than sediments aligns with previous studies (Wang et al., 2025; Jiang et al., 2018), which report a greater diversity and abundance of ARGs in surface water due to its dynamic nature and continuous input from anthropogenic sources. This supports the notion that water bodies act as conduits for ARG dissemination, facilitating their transport and potential HGT among microbial communities. However, sediments may serve as long-term reservoirs of ARGs, contributing to their persistence in aquatic environments.

The predominance of *sul-2, sul-1, tet(A),* and *tet(Q)* in water samples, particularly at Site 2, emphasizes the influence of specific environmental conditions and pollution sources. Sulfonamide and tetracycline resistance genes are commonly associated with agricultural and urban effluents in sediments, which may explain their elevated levels in aquatic ecosystems such as the Virilla Watershed. Interestingly, the enrichment of less abundant genes such as bla_{TEM}, *ermB,* and *qnrS* in water samples at the polluted Site 2 suggests a complex interplay of factors, including selective pressures from antibiotic residues and the presence of mobile genetic elements (MGEs) that facilitate HGT. The increased presence of *intI-1* at the most contaminated site (Site 3) further supports the role of MGEs in ARG propagation, as integrons are key players in capturing and disseminating ARGs.

Environmental factors such as temperature, pH, and dissolved oxygen (DO) play a crucial role in shaping microbial communities and influencing ARG distribution in aquatic environments (Deng et al., 2023; Niu et al., 2024; Yang et al., 2019; Yu et al., 2023). Temperature and nutrient availability, for example, directly impact HGT by affecting the prokaryotic membrane. Low temperatures (9–15 °C) negatively impact bacterial conjugation by reducing membrane fluidity, cellular metabolism, and pilus formation (Pallares-Vega et al., 2021). Similarly, nutrient scarcity significantly reduces plasmid transfer, likely due to restricted bacterial metabolism and diminished cell-to-cell interactions (Pallares-Vega et al., 2021). Previous studies indicate that elevated temperatures favor ARG enrichment and the proliferation of multidrug resistance genes, as observed in riverine and wetland ecosystems (Yang et al., 2019; Yu et al., 2023). Additionally, pH fluctuations can influence microbial community stability and ARG persistence by altering bacterial metabolism and DNA degradation dynamics (Niu et al., 2024). DO depletion, often associated with organic pollution, has been linked to shifts in microbial community composition and increased ARG mobility (Deng et al., 2023). Our study observed distinct variations in these physicochemical parameters along the pollution gradient, with higher temperatures and lower DO levels recorded at the most polluted sites. While these patterns align with previous research highlighting the role of environmental conditions in ARG persistence and dissemination, our findings emphasize that anthropogenic pollution is a key driver shaping these dynamics, acting as a selective force that accelerates antibiotic resistance evolution.

Recent studies further support the role of pollution as a key driver of ARG proliferation. Wang et al. (2025) demonstrated that ARG abundance increases at polluted sites, particularly in water samples, reinforcing our observations. Additionally, diverse anthropogenic activities, including urbanization, dam construction, and agricultural runoff, have significantly impacted ARG distribution (Liu et al., 2021). These activities promote conditions that favor ARG enrichment and dissemination, increasing their frequency in river systems. Urbanization, for instance, contributes to ARG diversity due to the influx of antibiotics and other pollutants from human activities such as wastewater discharge and agricultural runoff (Liu et al., 2021). Together, these findings strengthen the argument that urban pollution not only selects for ARG enrichment but also alters microbial community composition in ways that facilitate HGT.

The significant enrichment of ARGs and other cargo genes in the plasmidome along the contamination gradient, particularly in the water column, suggests that water samples may serve as hotspots for plasmid-mediated dissemination of resistance genes. This is likely due to the high mobility of genetic material in the water column compared to sediment, where transport is more constrained. Our findings align with previous research indicating that bacteria exhibit higher transport rates in the water column than in sediment communities. However, although this study did not quantify the horizontal transfer rate of MGEs, assessing this process is crucial for understanding the speed and extent of gene dissemination, which is key to addressing the spread of antibiotic resistance in microbial communities.

The findings on the dynamic evolution of plasmids in polluted sites reveal significant insights into the role of compound plasmids in gene transfer, particularly concerning resistance and metabolic traits. Our results align with previous research (Yu et al., 2024; Ciuchcinski, K. et al., 2024), underlining plasmids as crucial vectors for HGT in stressful environments. The formation of compound plasmids, as observed in this study, suggests that plasmid recombination is an adaptive mechanism in aquatic environments with high genetic mobility. This conclusion is supported by findings demonstrating how environmental factors drive the evolution of complex plasmid structures capable of harboring diverse genetic traits. Additionally, the observed site- and sample-type specificity of cargo genes aligns with previous findings, indicating that local microbial communities influence the structure and function of the plasmid metagenome.

Environmental pollution plays a crucial role in shaping microbial genetic landscapes. Polluted environments act as hotspots for the evolution and dissemination of resistance genes, posing significant challenges to public health and environmental management. Enriching plasmids with resistance and metabolic traits in polluted sites highlights the importance of understanding plasmids' ecological and evolutionary dynamics in response to environmental contamination.

Additionally, integrating qPCR and metagenomic approaches provides complementary insights into the environmental resistome. While qPCR enables highly sensitive and specific quantification of selected ARGs, metagenomics offers a broader and untargeted perspective on resistance gene diversity (Ferreira et al., 2023). In this study, metagenomic analyses captured a wide array of ARGs across pollution gradients, reinforcing that metagenomics is a powerful tool for surveying ARG diversity on a large scale (Liu et al., 2019). However, as reported in comparative studies, qPCR exhibited higher accuracy in quantifying ARG abundance, particularly for genes present at low concentrations in environmental matrices (Ferreira et al., 2023).

These results emphasize the need for a deeper understanding of the interplay between environmental conditions and ARG dissemination. Future research should aim to clarify the relative contributions of physicochemical factors and pollution-related stressors to antibiotic resistance trends in aquatic ecosystems. This knowledge is essential for predicting environmental drivers of ARG distribution and developing effective mitigation strategies. To achieve this, future studies should increase the number of physicochemical measurements, expand sampling sites, and conduct sampling across different time points or seasons. A multivariate statistical approach could further clarify the relative influence of each factor, including anthropogenic pollution, and provide insights into whether specific environmental parameters selectively influence plasmid-associated ARG fractions. This would improve our understanding of how ARGs persist and spread in aquatic environments.

4.5. Public health risks and the importance of monitoring environmental resistomes

Most detected ARGs confer resistance to aminoglycosides, beta-lactams, MLS (macrolide, lincosamide, streptogramin), and sulfonamides, with higher concentrations observed at the most contaminated

sites. The variety of ARG families identified also indicates the presence of complex contamination sources in the river, while the occurrence of the integrase class 1 gene (*intI-1*) in polluted sites can be related to enhanced recombination capacity (Singh et al., 2022; Zhuang et al., 2021). Despite the limitation in the sample size used in this study, this finding aligns with previous studies on ARGs detected in the surface water of the Virilla River Watershed, where this study was also conducted (Mendoza-Guido et al., 2024; Morales-Mora et al., 2025; Ramírez-Morales et al., 2021; Wilkinson et al., 2022). Given that more than half of the country's population resides in the GAM (Herrera-Murillo et al., 2019; Mena-Rivera et al., 2018), the area analyzed likely reflects the antibiotics usage. According to WHO reports, penicillin and other beta-lactams are among the most consumed antibiotic groups in Costa Rica, followed by other antimicrobials such as nitrofurantoin, MLS, tetracyclines, and sulfonamides (WHO, 2018; Zavaleta et al., 2023).

Notably, six mercury resistance genes were found more frequently in plasmids from the sediment samples of the most contaminated sites. Sediments can serve as reservoirs for these metal-resistance genes, potentially aiding their persistence and rapid dissemination in the environment. Studies have documented the presence of heavy metals, including mercury, in various regions of Costa Rica. For instance, mercury was detected in 61.5 % of sites' surface water sampled nationwide, often exceeding national regulatory limits (Quirós-Bustos et al., 2022). Furthermore, mining activities outside the GAM, such as in Tilarán, have been identified as significant sources of metal contamination, with leachates from abandoned sites contributing mercury and other metals to surrounding freshwater ecosystems. (Rojas-Conejo et al., 2021). In the Virilla River case, potential mercury sources in sediments include industrial discharges, urban runoff, and historical agricultural practices. These heavy metals can indirectly promote the proliferation of antibiotic-resistant bacteria by co-selecting ARGs (Di Cesare et al., 2022), thereby increasing human and environmental health risks.

We showed a considerable increase in ARGs, including those conferring resistance to aminoglycosides, beta-lactams, MLS, and sulfonamides. These were accompanied by increased fecal contaminants reflecting previously observed patterns in polluted environments (Karkman et al., 2019). This is accompanied by increased plasmid diversity and dynamics in their genetic makeup, with individual backbones of plasmids gaining genes with increasing pollution. Our results underscore the importance of distinguishing ARG dynamics on plasmids and chromosomes in environmental monitoring to predict regional resistance situations and to adapt antibiotic treatments. Characterizing the environmental resistome, particularly identifying the already mobilized and thus more easily transferrable fraction, is important to understand the role of the environment as a source for new resistance factors.

5. Conclusions

Pollution increases plasmidome diversity, promoting the spread of ARGs and altering microbial communities, especially in water. Studying both water and sediments is essential to understanding ARG dissemination.

Plasmids are key drivers of microbial adaptation, accelerating ARG acquisition under pollution stress. Their diversity makes them useful biomarkers for monitoring contamination and public health risks.

Environmental conditions influence ARG persistence and mobility, reinforcing the role of pollution as a selective force in resistance evolution.

Tracking plasmids improves AMR surveillance and helps differentiate between chromosomal and plasmid-borne ARGs, guiding better pollution management strategies.

Data availability statement

Supplemental material includes tables with microbiological and physicochemical data, primers, gBlocks sequences, and figures based on this study's results.

All metagenomic analysis software ran on the computer clusters SAGA (Sigma2, Norwegian Research Infrastructure Services) and Kabré (CeNAT Costa Rica), where SLURM 23.02.7 (76) jobs were written by 'metagenomix' (https://github.com/FranckLejzerowicz/metagenomix). Fully documented post-processing scripts and Python jupyter notebooks, as well as configuration files (databases, pipeline, modules, non-default parameters) to generate the commands used for all software and to replicate analyses are openly available at https://gitlab.com/alper1976/marmip/-/tree/main/keilor/papers/virilla_river/metagenomes.

The paired Fastq files containing the raw sequencing data, sample metadata, and experimental preparation information have been deposited to the European Nucleotide Archive under study accession number ERP163439 (runs accessions ERR13584077-ERR13584094).

CRediT authorship contribution statement

Kenia Barrantes-Jiménez: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Funding acquisition, Formal analysis. **Franck Lejzerowicz:** Writing – review & editing, Writing – original draft, Software, Methodology, Formal analysis, Data curation. **Tam Tran:** Writing – review & editing, Methodology. **Melany Calderón-Osorno:** Writing – review & editing, Methodology. **Luis Rivera-Montero:** Methodology, Investigation. **César Rodríguez-Sánchez:** Writing – review & editing, Investigation. **Odd-Gunnar Wikmark:** Investigation. **Alexander Eiler:** Writing – review & editing, Investigation. **Hans-Peter Grossart:** Writing – review & editing, Supervision, Investigation. **María Arias-Andrés:** Writing – review & editing, Visualization, Supervision, Methodology, Investigation, Funding acquisition, Conceptualization. **Keilor Rojas-Jiménez:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was funded by the Consejo Nacional de Rectores (CONARE), Costa Rica, the Vicerrectoría de Investigación of the Universidad de Costa Rica (projects C1455 and C2650; Rojas-Jiménez K, PI). The Universidad Nacional, Costa Rica (Project SIA 0483-21, Arias-Andrés M, PI), as part of the international MarMib project: Impact of Marine Microplastic-Associated Biofilms on Environmental and Human Health (Wikmark, O-G, PI) (https://prosjektbanken.forskningsradet.no/en/project/FORISS/315812). Additional funding was provided by the National Institutes of Health Fogarty International Center (grant number D43TW011403) for the project entitled '*International Training Program in Environmental Health over the Lifespan*' (Claudio L & van Wendel de Joode B, PIs), a grant awarded to the Icahn School of Medicine at Mount Sinai and Universidad Nacional, Costa Rica. The authors also extend their gratitude to Eddy Gómez from the Research Center for Marine and Limnological Sciences (CIMAR), UCR, for his support with field sampling and physicochemical analyses, as well as to Bradd Mendoza-Guido from the Health Research Institute (INISA-UCR) for his assistance in reviewing sections of the manuscript.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.watres.2025.123553.

References

Alonso-del Valle, A., León-Sampedro, R., Rodríguez-Beltrán, J., DelaFuente, J., Hernández-García, M., Ruiz-Garbajosa, P., Cantón, R., Peña-Miller, R., San Millán, A., 2021. Variability of plasmid fitness effects contributes to plasmid persistence in bacterial communities. Nat. Commun. 12 (1). https://doi.org/10.1038/s41467-021-22849-y.

American Public Health Association, American Water Works Association, & Water Environment Federation, 2017. In: Baird, R.B., Eaton, A.D., Rice, E.W. (Eds.), Standard Methods for the Examination of Water and Wastewater, 23rd ed. American Public Health Association.

Amon, J., Titgemeyer, F., Burkovski, A., 2009. A genomic view on nitrogen metabolism and nitrogen control in mycobacteria. J. Mol. Microbiol. Biotechnol. 17 (1), 20–29. https://doi.org/10.1159/000159195.

Arango-Argoty, G., Garner, E., Pruden, A., Heath, L.S., Vikesland, P., Zhang, L., 2018. DeepARG: a deep learning approach for predicting antibiotic resistance genes from metagenomic data. Microbiome 6 (1). https://doi.org/10.1186/s40168-018-0401-z.

Arias-Andres, M., Kettner, M.T., Miki, T., Grossart, H.P., 2018. Microplastics: new substrates for heterotrophic activity contribute to altering organic matter cycles in aquatic ecosystems. Sci. Total Environ. 635, 1152–1159. https://doi.org/10.1016/j.scitotenv.2018.04.199.

Bharat, A., Petkau, A., Avery, B.P., Chen, J., Folster, J., Carson, C.A., Kearney, A., Nadon, C., Mabon, P., Thiessen, J., Alexander, D.C., Allen, V., El Bailey, S., Bekal, S., German, G.J., Haldane, D., Hoang, L., Chui, L., Minion, J., Mulvey, M.R., 2022. Correlation between phenotypic and in silico detection of antimicrobial resistance in Salmonella enterica in Canada using Staramr. Microorganisms 10 (2). https://doi.org/10.3390/microorganisms10020292.

Barrantes, K., Chacón, L.M., Morales, E., Ramírez-Carvajal, L., 2020. Draft genome sequence of an *Escherichia coli* strain harboring bla CTX-M-115, bla CMY-2, aminoglycoside, tetracycline, and sulfonamide resistance genes, isolated from a Costa Rican wastewater treatment plant. Microbiol. Resour. Announc. 9 (1), 1–3. https://doi.org/10.1128/MRA.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J.E., Bittinger, K., Brejnrod, A., Brislawn, C.J., Brown, C.T., Callahan, B.J., Caraballo-Rodríguez, A.M., Chase, J., Caporaso, J.G., 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat. Biotechnol. 37 (8), 852–857. https://doi.org/10.1038/s41587-019-0209-9.

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., Neumann, B., Werner, G., Bender, J.K., Stingl, K., Nguyen, M., Coppens, J., Xavier, B. B., Aarestrup, F.M., 2020. ResFinder 4.0 for predictions of phenotypes from genotypes. J. Antimicrob. Chemother. 75 (12), 3491–3500. https://doi.org/10.1093/jac/dkaa345.

Bottery, M.J., 2022. Ecological dynamics of plasmid transfer and persistence in microbial communities. Curr. Opin. Microbiol. 68. https://doi.org/10.1016/j.mib.2022.102152.

Brown, R.M., McClelland, N.I., Deininger, R.A., O'Connor, M.F, 1972. A water quality index-crashing the psychological barrier. Indicators of Enviromental Quality.

Cantalapiedra, C.P., Herṇandez-Plaza, A., Letunic, I., Bork, P., Huerta-Cepas, J., 2021. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. Mol. Biol. Evol. 38 (12), 5825–5829. https://doi.org/10.1093/molbev/msab293.

Carattoli, A., Zankari, E., Garciá-Fernández, A., Larsen, M.V., Lund, O., Villa, L., Aarestrup, F.M., Hasman, H., 2014. In Silico detection and typing of plasmids using plasmidfinder and plasmid multilocus sequence typing. Antimicrob. Agents Chemother 58 (7), 3895–3903. https://doi.org/10.1128/AAC.02412-14.

Chen, S., Zhou, Y., Chen, Y., Gu, J., 2018. Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34 (17), i884–i890. https://doi.org/10.1093/bioinformatics/bty560.

Ciuchcinski, K., Stokke, R., Steen, I.H., Dziewit, L., 2024. Landscape of the metaplasmidome of deep-sea hydrothermal vents located at Arctic Mid-Ocean Ridges in the Norwegian–Greenland Sea: ecological insights from comparative analysis of plasmid identification tools. FEMS. Microbiol. Ecol. (10), 100. https://doi.org/10.1093/femsec/fiae124.

Conway, J.R., Lex, A., Gehlenborg, N., 2017. UpSetR: an R package for the visualization of intersecting sets and their properties. Bioinformatics 33 (18), 2938–2940. https://doi.org/10.1093/bioinformatics/btx364.

Czatzkowska, M., Wolak, I., Harnisz, M., Korzeniewska, E., 2022. Impact of anthropogenic activities on the dissemination of ARGs in the environment—a review. Int. J. Environ. Res. Public Health 19 (19). https://doi.org/10.3390/ijerph191912853.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., 2021. Twelve years of SAMtools and BCFtools. Gigascience (2), 10. https://doi.org/10.1093/gigascience/giab008.

Datt, S., Gregorio, D., 2024. hAMRonization: enhancing antimicrobial resistance prediction using the PHA4GE AMR detection specification and tooling. bioRxiv. March, 1–23. https://doi.org/10.1101/2024.03.07.583950.

de la Cruz, E.D.la, Garcia, F., Molina, A., 2014. Hazard prioritization and risk characterization of antibiotics in an irrigated Costa Rican region used for intensive crop, livestock and aquaculture farming. J. Environ. Biol. 35 (Special issue), 85–98.

de Lipthay, J.R., Rasmussen, L.D., Oregaard, G., Simonsen, K., Bahl, M.I., Kroer, N., Sørensen, S.J., 2008. Acclimation of subsurface microbial communities to mercury. FEMS Microbiol. Ecol. 65 (1), 145–155. https://doi.org/10.1111/j.1574-6941.2008.00501.x.

Deng, Y., Jiang, J., Huang, Y., Cheng, C., Lin, Z., Liu, G., Guo, Z., Feng, J., 2023. Hypoxia triggers the proliferation of antibiotic resistance genes in a marine aquaculture system. Sci. Total Environ. 859. https://doi.org/10.1016/j.scitotenv.2022.160305.

Di Cesare, A., Sabatino, R., Yang, Y., Brambilla, D., Li, P., Fontaneto, D., Eckert, E.M., Corno, G., 2022. Contribution of plasmidome, metal resistome and integrases to the persistence of the antibiotic resistome in aquatic environments. Environ. Pollut. 297. https://doi.org/10.1016/j.envpol.2021.118774.

Ebmeyer, S., Kristiansson, E., Larsson, D.G.J., 2021. A framework for identifying the recent origins of mobile antibiotic resistance genes. Commun. Biol. 4 (1), 1–10. https://doi.org/10.1038/s42003-020-01545-5.

Ferreira, C., Otani, S., Aarestrup, F.M., Manaia, C.M., 2023. Quantitative PCR versus metagenomics for monitoring antibiotic resistance genes: balancing high sensitivity and broad coverage. FEMS Microbes 4. https://doi.org/10.1093/femsmc/xtad008.

Feldgarden, M., Brover, V., Gonzalez-Escalona, N., Frye, J.G., Haendiges, J., Haft, D.H., Hoffmann, M., Pettengill, J.B., Prasad, A.B., Tillman, G.E., Tyson, G.H., Klimke, W., 2021. AMRFinderPlus and the Reference Gene Catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. Sci. Rep. 11 (1). https://doi.org/10.1038/s41598-021-91456-0.

Gillings, M.R., 2017. Lateral gene transfer, bacterial genome evolution, and the Anthropocene. Ann. N. Y. Acad. Sci. 1389 (1), 20–36. https://doi.org/10.1111/nyas.13213.

Herrera-Murillo, J., Anchía-Leitón, D., Rojas-Marín, J.F., Mora-Campos, D., Gamboa-Jiménez, A., Chaves-Villalobos, M., 2019. Influence of land use patterns on the quality of surface waters in the Virilla River sub-basin, Costa Rica. Rev. Geogr. Am. Central 4 (61E), 11. https://doi.org/10.15359/rgac.61-4.1.

Huerta-Cepas, J., Forslund, K., Coelho, L.P., Szklarczyk, D., Jensen, L.J., Von Mering, C., Bork, P., 2017. Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. Mol. Biol. Evol. 34 (8), 2115–2122. https://doi.org/10.1093/molbev/msx148.

Hyatt, D., Chen, G.-L., Locascio, P.F., Land, M.L., Larimer, F.W., Hauser, L.J., 2010. Prodigal: Prokaryotic Gene Recognition and Translation Initiation Site Identification. http://www.biomedcentral.com/1471-2105/11/119.

Ichwana, I., Syahrul, S., Nelly, W., 2016. Water quality index by using national sanitation foundation-Water Quality Index (NSF-WQI) Method at Krueng Tamiang Aceh. In: Proceeding of the First International Conference on Technology, Innovation and Society, pp. 110–117. https://doi.org/10.21063/ICTIS.2016.1019.

Jiang, C., Chen, H., Grossart, H.-P., Zhang, Q., Stoks, R., Zhao, Y., Ju, F., Liu, W., Yang, Y., 2022. Frequency of occurrence and habitat selection shape the spatial variation in the antibiotic resistome in riverine ecosystems in eastern China. Environ. Microbiome 17 (1), 53. https://doi.org/10.1186/s40793-022-00447-9.

Jiang, H., Zhou, R., Zhang, M., Cheng, Z., Li, J., Zhang, G., Chen, B., Zou, S., Yang, Y., 2018. Exploring the differences of antibiotic resistance genes profiles between river surface water and sediments using metagenomic approach. Ecotoxicol. Environ. Saf. 161 (February), 64–69. https://doi.org/10.1016/j.ecoenv.2018.05.044.

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), 1–8. https://doi.org/10.1038/s41467-018-07992-3.

Karlicki, M., Antonowicz, S., Karnkowska, A., 2022. Tiara: deep learning-based classification system for eukaryotic sequences. Bioinformatics 38 (2), 344–350. https://doi.org/10.1093/bioinformatics/btab672.

Kothari, A., Wu, Y.W., Chandonia, J.M., Charrier, M., Rajeev, L., Rocha, A.M., Joyner, D. C., Hazen, T.C., Singer, S.W., Mukhopadhyay, A., 2019. Large circular plasmids from groundwater plasmidomes span multiple incompatibility groups and are enriched in multimetal resistance genes. mBio (1), 10. https://doi.org/10.1128/mBio.02899-18.

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9 (4), 357–359. https://doi.org/10.1038/nmeth.1923.

Larsson, D.G.J., Flach, C.F., 2022. Antibiotic resistance in the environment. Nat. Rev. Microbiol. 20 (5), 257–269. https://doi.org/10.1038/s41579-021-00649-x.

Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R., Pfister, H., 2014. UpSet: visualization of intersecting sets. IEEE Trans. Vis. Comput. Graph. 20 (12), 1983–1992. https://doi.org/10.1109/TVCG.2014.2346248.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25 (16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

Liu, S., Wang, P., Wang, C., Wang, X., Chen, J., 2021. Anthropogenic disturbances on antibiotic resistome along the Yarlung Tsangpo River on the Tibetan Plateau: ecological dissemination mechanisms of antibiotic resistance genes to bacterial pathogens. Water Res. 202. https://doi.org/10.1016/j.watres.2021.117447.

Liu, X., Xiao, P., Guo, Y., Liu, L., Yang, J., 2019. The impacts of different high-throughput profiling approaches on the understanding of bacterial antibiotic resistance genes in a freshwater reservoir. Sci. Total Environ. 693. https://doi.org/10.1016/j.scitotenv.2019.133585.

Lu, J., Breitwieser, F.P., Thielen, P., Salzberg, S.L., 2017. Bracken: estimating species abundance in metagenomics data. PeerJ Comput. Sci. 2017 (1). https://doi.org/10.7717/peerj-cs.104.

Magoč, T., Salzberg, S.L., 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27 (21), 2957–2963. https://doi.org/10.1093/bioinformatics/btr507.

Marselina, M., Wibowo, F., Mushfiroh, A., 2022. Water quality index assessment methods for surface water: a case study of the Citarum River in Indonesia. Heliyon 8 (7). https://doi.org/10.1016/j.heliyon.2022.e09848.

Mena-Rivera, L., Vásquez-Bolaños, O., Gómez-Castro, C., Fonseca-Sánchez, A., Rodríguez-Rodríguez, A., Sánchez-Gutiérrez, R., 2018. Ecosystemic assessment of surface water quality in the Virilla River: towards sanitation processes in Costa Rica. Water 10 (7), 1–16. https://doi.org/10.3390/w10070845.

Mendoza-Guido, B., Barrantes, K., Rodríguez, C., Rojas-Jimenez, K., Arias-Andres, M., 2024. The impact of urban pollution on plasmid-mediated resistance acquisition in enterobacteria from a tropical river. Antibiotics 13 (11). https://doi.org/10.3390/antibiotics13111089.

Mora-Aparicio, C., Alfaro-Chinchilla, C., Pérez-Molina, J.P., Vega-Guzmán, I., 2022. Environmental contribution of Los Tajos wastewater treatment plant in the removal of physicochemical and microbiological pollutants. Uniciencia 36 (1), 1–17. https://doi.org/10.15359/ru.36-1.33.

Morales-Mora, E., Rivera-Montero, L., Montiel-Mora, J.R., Barrantes-Jiménez, K., Chacón-Jiménez, L., 2025. Assessing microbial risks of *Escherichia coli*: a spatial and temporal study of virulence and resistance genes in surface water in resource-limited regions. Sci. Total Environ. 958, 178044. https://doi.org/10.1016/j.scitotenv.2024.178044.

Neethu, C.S., Saravanakumar, C., Purvaja, R., Robin, R.S., Ramesh, R., 2022. Arsenic resistance and horizontal gene transfer are associated with carbon and nitrogen enrichment in bacteria. Environ. Pollut. 311. https://doi.org/10.1016/j.envpol.2022.119937.

Newbury, A., Dawson, B., Kl€ Umper, U., Hesse, E., Castledine, M., Fontaine, C., Buckling, A., & Sanders, D. (2022). *Fitness Effects of Plasmids Shape the Structure of Bacteria-Plasmid Interaction Networks*. https://doi.org/10.1073/pnas.

Nishimura, Y., Watai, H., Honda, T., Mihara, T., Omae, K., Roux, S., Blanc-Mathieu, R., Yamamoto, K., Hingamp, P., Sako, Y., Sullivan, M.B., Goto, S., Ogata, H., Yoshida, T., 2017. Environmental viral genomes shed new light on virus-host interactions in the ocean. mSphere 2 (2). https://doi.org/10.1128/msphere.00359-16.

Niu, C., Wang, B., Wang, Z., Zhang, H., 2024. Effect of pH on antibiotic resistance genes removal and bacterial nucleotides metabolism function in the wastewater by the combined ferrate and sulfite treatment. Chem. Eng. J. 480. https://doi.org/10.1016/j.cej.2023.148042.

Numberger, D., Zoccarato, L., Woodhouse, J., Ganzert, L., Sauer, S., Márquez, J.R.G., Domisch, S., Grossart, H.P., Greenwood, A.D., 2022. Urbanization promotes specific bacteria in freshwater microbiomes including potential pathogens. Sci. Total Environ. 845. https://doi.org/10.1016/j.scitotenv.2022.157321.

Nurk, S., Meleshko, D., Korobeynikov, A., Pevzner, P.A., 2017. MetaSPAdes: a new versatile metagenomic assembler. Genome Res. 27 (5), 824–834. https://doi.org/10.1101/gr.213959.116.

Pallares-Vega, R., Macedo, G., Brouwer, M.S.M., Hernandez Leal, L., van der Maas, P., van Loosdrecht, M.C.M., Weissbrodt, D.G., Heederik, D., Mevius, D., Schmitt, H., 2021. Temperature and nutrient limitations decrease transfer of conjugative IncP-1 plasmid pKJK5 to wild *Escherichia coli* strains. Front. Microbiol. 12. https://doi.org/10.3389/fmicb.2021.656250.

Parks, D.H., Chuvochina, M., Waite, D.W., Rinke, C., Skarshewski, A., Chaumeil, P.A., Hugenholtz, P., 2018. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat. Biotechnol. 36 (10), 996. https://doi.org/10.1038/nbt.4229.

Partridge, S.R., Kwong, S.M., Firth, N., Jensen, S.O., 2018. Mobile genetic elements associated with antimicrobial resistance. Clin. Microbiol. Rev. 31 (4). https://doi.org/10.1128/CMR.00088-17.

Pérez Gómez, G., Alvarado García, V., Rodríguez Rodríguez, J.A., Herrera, F., Sánchez Gutiérrez, R., 2021. Calidad fisicoquímica y microbiológica del agua superficial del río Grande de Tárcoles, Costa Rica: un enfoque ecológico. UNED Res. J. 13 (1), 17. https://doi.org/10.22458/urj.v13i1.3148.

Petersen, J., Vollmers, J., Ringel, V., Brinkmann, H., Ellebrandt-Sperling, C., Spröer, C., Howat, A.M., Murrell, J.C., Kaster, A.K., 2019. A marine plasmid hitchhiking vast phylogenetic and geographic distances. Proc. Natl. Acad. Sci. USA 116 (41), 20568–20573. https://doi.org/10.1073/pnas.1905878116.

Pradier, L., Tissot, T., Fiston-Lavier, A.S., Bedhomme, S., 2021. PlasForest: a homology-based random forest classifier for plasmid detection in genomic datasets. BMC Bioinformatics 22 (1), 349. https://doi.org/10.1186/s12859-021-04270-w.

Quirós-Bustos, N., Robles-Chaves, D., Caballero-Chavarría, A., Calvo-Brenes, G., 2022. Heavy metal content in various rivers of Costa Rica. [Contenido de metales pesados en varios ríos de Costa Rica]. Rev. Tecnol. Marcha. https://doi.org/10.18845/tm.v35i2.5532.

Ramírez-Morales, D., Masís-Mora, M., Montiel-Mora, J.R., Cambronero-Heinrichs, J.C., Pérez-Rojas, G., Tormo-Budowski, R., Méndez-Rivera, M., Briceño-Guevara, S., Gutiérrez-Quirós, J.A., Arias-Mora, V., Brenes-Alfaro, L., Beita-Sandí, W., Rodríguez-Rodríguez, C.E., 2021. Multi-residue analysis of pharmaceuticals in water samples by liquid chromatography- mass spectrometry: quality assessment and application to the risk assessment of urban-influenced surface waters in a metropolitan area of Central America. Process Saf. Environ. Protect. 153, 289–300. https://doi.org/10.1016/j.psep.2021.07.025.

Rodríguez-Beltrán, J., DelaFuente, J., León-Sampedro, R., MacLean, R.C., San Millán, Á., 2021. Beyond horizontal gene transfer: the role of plasmids in bacterial evolution. Nat. Rev. Microbiol. 19 (6), 347–359. https://doi.org/10.1038/s41579-020-00497-1.

Rojas-Conejo, J., Picado Pavón, F., Suárez Serrano, A., Van Gestel, C.A.M., Golcher Benavides, C., Durán Sanabria, G., 2021. Mining environmental liabilities: a potential source of metal contamination for freshwater ecosystems in Costa Rica. Rev. Geogr. Am. Central 1 (68), 333–356. https://doi.org/10.15359/rgac.68-1.12.

Rojas-Villalta, D., Calderón-Osorno, M., Barrantes-Jiménez, K., Arias-Andres, M., Rojas-Jiménez, K., 2023. Benchmarking AI-based plasmid annotation tools for antibiotic resistance genes mining from metagenome of the Virilla River, Costa Rica. bioRxiv. https://doi.org/10.1101/2023.08.24.554652.

Schmitt, A.O., Herzel, H., 1997. Estimating the entropy of DNA sequences introduction: order and disorder of sequences. J. Theor. Biol. 1888, 369–377.

Schloss, P.D., 2024. Rarefaction is currently the best approach to control for uneven sequencing effort in amplicon sequence analyses. mSphere 9 (2). https://doi.org/10.1128/msphere.00354-23.

Schwengers, O., Barth, P., Falgenhauer, L., Hain, T., Chakraborty, T., Goesmann, A., 2020. Platon: identification and characterization of bacterial plasmid contigs in short-read draft assemblies exploiting protein sequence-based replicon distribution scores. Microb. Genom. 6 (10), 1–12. https://doi.org/10.1099/mgen.0.000398.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T., 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13 (11), 2498–2504. https://doi.org/10.1101/gr.1239303.

Sherry, N.L., Horan, K.A., Ballard, S.A., Gonçalves da Silva, A., Gorrie, C.L., Schultz, M.B., Stevens, K., Valcanis, M., Sait, M.L., Stinear, T.P., Howden, B.P., Seemann, T., 2023. An ISO-certified genomics workflow for identification and surveillance of antimicrobial resistance. Nat. Commun. 14 (1). https://doi.org/10.1038/s41467-022-35713-4.

Shintani, M., Sanchez, Z.K., Kimbara, K., 2015. Genomics of microbial plasmids: classification and identification based on replication and transfer systems and host taxonomy. Front. Microbiol. 6 (MAR). https://doi.org/10.3389/fmicb.2015.00242.

Singh, A.K., Kaur, R., Verma, S., Singh, S., 2022. Antimicrobials and antibiotic resistance genes in water bodies: pollution, risk, and control. Front. Environ. Sci. 10. https://doi.org/10.3389/fenvs.2022.830861.

Slater, F.R., Bailey, M.J., Tett, A.J., Turner, S.L., 2008. Progress towards understanding the fate of plasmids in bacterial communities. FEMS. Microbiol. Ecol. 66 (1), 3–13. https://doi.org/10.1111/j.1574-6941.2008.00505.x.

Smalla, K., Cook, K., Djordjevic, S.P., Klümper, U., Gillings, M., 2018. Environmental dimensions of antibiotic resistance: assessment of basic science gaps. FEMS. Microbiol. Ecol. 94 (12), 1–6. https://doi.org/10.1093/femsec/fiy195.

Smillie, C., Garcillán-Barcia, M.P., Francia, M.V., Rocha, E.P.C., de la Cruz, F., 2010. Mobility of plasmids. Microbiol. Mol. Biol. Rev. 74 (3), 434–452. https://doi.org/10.1128/mmbr.00020-10.

Spongberg, A.L., Witter, J.D., Acuña, J., Vargas, J., Murillo, M., Umaña, G., Gómez, E., Perez, G., 2011. Reconnaissance of selected PPCP compounds in Costa Rican surface waters. Water Res. 45 (20), 6709–6717. https://doi.org/10.1016/j.watres.2011.10.004.

Stalder, T., Cornwell, B., Lacroix, J., Kohler, B., Dixon, S., Yano, H., Kerr, B., Forney, L.J., Top, E.M., 2020. Evolving populations in biofilms contain more persistent plasmids. Mol. Biol. Evol. 37 (6), 1563–1576. https://doi.org/10.1093/molbev/msaa024.

Stockdale, S.R., Harrington, R.S., Shkoporov, A.N., Khokhlova, E.v., Daly, K.M., McDonnell, S.A., O'Reagan, O., Nolan, J.A., Sheehan, D., Lavelle, A., Draper, L.A., Shanahan, F., Ross, R.P., Hill, C., 2022. Metagenomic assembled plasmids of the human microbiome vary across disease cohorts. Sci. Rep. 12 (1). https://doi.org/10.1038/s41598-022-13313-y.

Stockdale, S.R., Hill, C., 2023. Incorporating plasmid biology and metagenomics into a holistic model of the human gut microbiome. Curr. Opin. Microbiol. 73. https://doi.org/10.1016/j.mib.2023.102307.

Suzuki, S., Pruden, A., Virta, M., Zhang, T., 2017. Antibiotic resistance in aquatic systems. Front. Microbiol. 8 (April). https://doi.org/10.3389/fmicb.2017.00014. Frontiers Research Topics.

Topp, E., Larsson, D.G.J., Miller, D.N., Van den Eede, C., Virta, M.P.J, 2018. Antimicrobial resistance and the environment: assessment of advances, gaps and recommendations for agriculture, aquaculture and pharmaceutical manufacturing. FEMS Microbiol. Ecol. 94 (3), 1–5. https://doi.org/10.1093/femsec/fix185.

Uddin, M.G., Nash, S., Olbert, A.I., 2021. A review of water quality index models and their use for assessing surface water quality. Ecol. Indicat. 122. https://doi.org/10.1016/j.ecolind.2020.107218.

Vale, F.F., Lehours, P., Yamaoka, Y., 2022. Editorial: the role of mobile genetic elements in bacterial evolution and their adaptability. Front. Microbiol. 13. https://doi.org/10.3389/fmicb.2022.849667.

van Passel, M.W.J., Bart, A., Luyf, A.C.M., van Kampen, A.H.C., van der Ende, A., 2006. Compositional discordance between prokaryotic plasmids and host chromosomes. BMC Genomics 7. https://doi.org/10.1186/1471-2164-7-26.

Venegas González, D.A., Morales Mora, E., Barrantes Jiménez, K., Gómez Ramírez, E., Fuentes-Schweizer, P., Irías Mata, A., 2023. Contaminación del agua del río Durazno, Costa Rica: más allá del índice holandés de calidad del agua. UNED Res. J. 15 (1), e4339. https://doi.org/10.22458/urj.v15i1.4339.

Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Vázquez-Baeza, Y., 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods 17 (3), 261–272. https://doi.org/10.1038/s41592-019-0686-2.

Wallace, J.C., Youngblood, J.E., Port, J.A., Cullen, A.C., Smith, M.N., Workman, T., Faustman, E.M., 2018. Variability in metagenomic samples from the Puget Sound: relationship to temporal and anthropogenic impacts. PLoS ONE 13 (2). https://doi.org/10.1371/journal.pone.0192412.

Wang, Y., Dagan, T., 2024. The evolution of antibiotic resistance islands occurs within the framework of plasmid lineages. Nat. Commun. 15 (1). https://doi.org/10.1038/s41467-024-48352-8.

Wang, G., Haenelt, S., Corrêa, F.B., da Rocha, U.N., Musat, F., Zhang, J., Müller, J.A., Musat, N., 2025. Riverine antibiotic resistome along an anthropogenic gradient. Front. Microbiol. 16. https://doi.org/10.3389/fmicb.2025.1516033.

WHO, 2018. WHO Report on Surveillance of Antibiotic Consumption: 2016-2018 Early Implementation. https://www.who.int/publications/i/item/9789241514880.

Wilkinson, J.L., Boxall, A.B.A., Kolpin, D.W., Leung, K.M.Y., Lai, R.W.S., Galbán-Malagón, C., Adell, A.D., Mondon, J., Metian, M., Marchant, R.A., Bouzas-

Monroy, A., Cuni-Sanchez, A., Coors, A., Carriquiriborde, P., Rojo, M., Gordon, C., Cara, M., Moermond, M., Luarte, T., Teta, C., 2022. Pharmaceutical pollution of the world's rivers. Proc. Natl. Acad. Sci. 119 (8). https://doi.org/10.1073/pnas.2113947119.

Wood, D.E., Lu, J., Langmead, B., 2019. Improved metagenomic analysis with Kraken 2. Genome Biol. 20 (1). https://doi.org/10.1186/s13059-019-1891-0.

Yang, Y., Liu, G., Ye, C., Liu, W., 2019. Bacterial community and climate change implication affected the diversity and abundance of antibiotic resistance genes in wetlands on the Qinghai-Tibetan Plateau. J. Hazard. Mater. 361, 283–293. https://doi.org/10.1016/j.jhazmat.2018.09.002.

Yoo, A.B., Jette, M.A., Grondona, M., 2003. LNCS 2862 - SLURM: Simple Linux Utility For Resource Management.

Yu, M.K., Fogarty, E.C., Eren, A.M., 2024. Diverse plasmid systems and their ecology across human gut metagenomes revealed by PlasX and MobMess. Nat. Microbiol. 9 (3), 830–847. https://doi.org/10.1038/s41564-024-01610-3.

Yu, Q., Han, Q., Shi, S., Sun, X., Wang, X., Wang, S., Yang, J., Su, W., Nan, Z., Li, H., 2023. Metagenomics reveals the response of antibiotic resistance genes to elevated temperature in the Yellow River. Sci. Total Environ. 859. https://doi.org/10.1016/j.scitotenv.2022.160324.

Zavaleta, E., Ferrara, F., Zovi, A., Díaz-Madriz, J.P., Fallas-Mora, A., Serrano-Arias, B., Valentino, F., Arguedas-Chacón, S., Langella, R., Trama, U., Nava, E., 2023. Antibiotic consumption in primary care in Costa Rica and Italy: a retrospective cross-country analysis. Cureus. https://doi.org/10.7759/cureus.41414.

Zhang, H., Yang, L., Li, Y., Wang, C., Zhang, W., Wang, L., Niu, L., 2022. Pollution gradients shape the co-occurrence networks and interactions of sedimentary bacterial communities in Taihu Lake, a shallow eutrophic lake. J. Environ. Manage 305. https://doi.org/10.1016/j.jenvman.2021.114380.

Zhuang, M., Achmon, Y., Cao, Y., Liang, X., Chen, L., Wang, H., Siame, B.A., Leung, K.Y., 2021. Distribution of antibiotic resistance genes in the environment. Environ. Pollut. 285. https://doi.org/10.1016/j.envpol.2021.117402.