



HAL
open science

The relationship between water quality and the microbial virulome and resistome in urban streams in Brazil

Eliana Guedes Stehling, João Pedro Rueda Furlan, Ralf Lopes, John Chodkowski, Nejc Stopnisek, Eduardo Angelino Savazzi, Ashley Shade

► To cite this version:

Eliana Guedes Stehling, João Pedro Rueda Furlan, Ralf Lopes, John Chodkowski, Nejc Stopnisek, et al.. The relationship between water quality and the microbial virulome and resistome in urban streams in Brazil. *Environmental Pollution*, 2024, 348, pp.12349. 10.1016/j.envpol.2024.123849 . hal-04590406

HAL Id: hal-04590406

<https://hal.science/hal-04590406v1>

Submitted on 28 May 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 **The relationship between water quality and the microbial virulome and**
2 **resistome in urban streams in Brazil**

3

4

5 Eliana Guedes Stehling^{a,*}, João Pedro Rueda Furlan^a, Ralf Lopes^a, John Chodkowski^b, Nejc
6 Stopnisek^{b,c}, Eduardo Angelino Savazzi^d, and Ashley Shade^e

7

8

9 ^a Department of Clinical Analyses, Toxicology and Food Science, School of Pharmaceutical
10 Sciences of Ribeirão Preto, University of São Paulo, Brazil.

11 ^b Department of Microbiology and Molecular Genetics, Michigan State University (MSU) –
12 East Lansing, United States.

13 ^c National Laboratory of Health, Environment and Food – Maribor, Slovenia.

14 ^d Environmental Company of São Paulo State, Ribeirão Preto, Brazil.

15 ^e Université Claude Bernard Lyon 1, Laboratoire d'Ecologie Microbienne, UMR CNRS 5557,
16 UMR INRAE 1418, VetAgro Sup, 69622 Villeurbanne, France.

17

18

19 *** Corresponding author:**

20 Prof. Eliana Guedes Stehling, Ph.D. (elianags@usp.br)

21 Mailing address: Av. do Café, s/n, Monte Alegre, Ribeirão Preto, 14040-903, Brazil.

22 Phone: + 55 16 3315 0285

23

24 **Abstract**

25

26 Urban streams that receive untreated domestic and hospital waste can transmit infectious
27 diseases and spread drug residues, including antimicrobials, which can then increase the
28 selection of antimicrobial-resistant bacteria. Here, water samples were collected from three
29 different urban streams in the state of São Paulo, Brazil, to relate their range of Water Quality
30 Indices (WQIs) to the diversity and composition of aquatic microbial taxa, virulence genes
31 (virulome), and antimicrobial resistance determinants (resistome), all assessed using untargeted
32 metagenome sequencing. There was a predominance of phyla Proteobacteria, Actinobacteria,
33 and Bacteroidetes in all samples, and *Pseudomonas* was the most abundant detected genus.
34 Virulence genes associated with motility, adherence, and secretion systems were highly
35 abundant and mainly associated with *Pseudomonas aeruginosa*. Furthermore, some
36 opportunistic pathogenic genera had negative correlations with WQI. Many clinically relevant
37 antimicrobial resistance genes (ARGs) and efflux pump-encoding genes that confer resistance
38 to critically important antimicrobials were detected. The highest relative abundances of ARGs
39 were β -lactams and macrolide-lincosamide-streptogramin. No statistically supported
40 relationship was detected between the abundance of virulome/resistome and collection
41 type/WQI. On the other hand, total solids were a weak predictor of gene abundance patterns.
42 These results provide insights into various microbial outcomes given urban stream quality and
43 point to its ecological complexity. In addition, this study suggests potential consequences for
44 human health as mediated by aquatic microbial communities responding to typical urban
45 outputs.

46

47 **Keywords:** Microbial diversity; Virulome; Resistome; Metagenomics; Water.

48

49 **1. Introduction**

50

51 Many streams and springs are used for public supply or development of economic
52 activities, including dams, rivers, lakes, and groundwater (Loucks & van Beek, 2017).
53 However, watersheds can become unsuitable for supply for various reasons, including lack of
54 sanitation infrastructure, erosion of rivers and streams, inadequate exploitation of water
55 resources, unregulated industrial activities, and deposition of domestic and hospital sewage.
56 These activities can result in poor water quality and expose the human population to disease
57 (Kumari et al., 2020; Majumder et al., 2021). Among the pollutants that affect urban public
58 water sources, domestic sewage, hospital effluents, heavy metals, organochlorine substances,
59 and derivatives stand out. The release of untreated domestic and hospital effluents into rivers
60 causes an increase in organic matter in the water and an imbalance in the ecosystem, which
61 increases some populations of microorganisms and eliminates others (Muniz et al., 2015). This
62 process can also lead to the transmission of diseases due to the presence of pathogenic bacteria
63 in untreated human waste. In addition, domestic sewage may also be contaminated by drug
64 residues, including antimicrobials, which affect aquatic environments across the planet (Ortúzar
65 et al., 2022).

66 It is hypothesized that variations of water quality indices (WQIs) may be related to the
67 increase, dissemination, and persistence of medically important bacteria and their antimicrobial
68 resistance genes (ARGs) in aquatic ecosystems as follows: i) pH affects bacterial populations,
69 including pathogenic bacteria and their antimicrobial resistance (AMR) profiles and influences
70 the prevalence and distribution of ARGs-producing bacterial strains (Li et al., 2015); ii)
71 increased nutrient levels enhance the proliferation of antimicrobial-resistant bacteria, as well as
72 the horizontal transfer of ARGs among microbial populations (Yang et al., 2014); iii) high
73 biological oxygen demand (BOD) levels favor conditions for microbial growth and persistence

74 (Pruden et al., 2013); iv) increased turbidity is correlated with higher bacterial loads and a
75 greater abundance of ARGs, as suspended particles can facilitate the transport and spread of
76 bacterial and xenogenetic pollutants (Guo et al., 2020); v) temperature fluctuations alter the
77 composition and activity of bacterial populations, influencing the prevalence and persistence of
78 AMR (Zhang et al., 2009); vi) elevated levels of total solids and total dissolved solids support
79 microbial growth and survival, contributing to increased bacterial densities and the
80 dissemination of AMR determinants (Stalder et al., 2019); vii) variations in dissolved oxygen
81 concentrations impact the composition and functional profiles of bacterial communities,
82 affecting the prevalence and persistence of AMR traits (Winker et al., 2020); and viii) elevated
83 fecal coliform levels are closely correlated with increased abundance and diversity of antibiotic-
84 resistant bacteria and ARGs in water, reflecting the introduction of resistant strains from human
85 and animal fecal sources (Jutkina et al., 2016).

86 In this context, this study aimed to evaluate and compare the microbial diversity,
87 virulome, and resistome of three important water sources historically monitored by the
88 Environmental Company of São Paulo State (CETESB) and having different water quality
89 levels. To compare these streams, we used untargeted metagenome sequencing to assess the
90 composition and diversity of functional genes for virulence factors and ARGs. Our objective
91 was to understand how water quality influences the maintenance of these metagenome
92 signatures, including the presence of pathogenic bacteria, virulence genes, ARGs, and efflux
93 pump-encoding genes in these critical water sources for a key urban-supportive watershed in
94 Brazil.

95

96 **2. Material and methods**

97

98 *2.1. Sample collection and indices*

99

100 Water samples were collected by the CETESB, which is responsible for establishing the
101 indices and environmental indicators of the aquatic environments of São Paulo, Brazil. Water
102 samples were collected from three different aquatic ecosystems in the state of São Paulo: São
103 Domingos Stream (1 site; agricultural area), Pardo River (2 sites; industrial area), and Tietê
104 River (5 sites; industrial area). These collection sites were selected because they represent a
105 range of different urban influences and water qualities, as classified by the WQIs, which can be
106 categorically defined as poor (20 to 36), regular (37 to 51), good (52 to 79), and excellent (80
107 to 100). The WQI was calculated for each site based on nine quality indicators (pH, total
108 nitrogen, total phosphorus, biochemical oxygen demand, turbidity, temperature, total residue,
109 dissolved oxygen, and fecal coliforms) (Supplementary Table S1). All parameters and indices
110 are available in CETESB publications and reports ([https://cetesb.sp.gov.br/praias/publicacoes-
111 relatorios/](https://cetesb.sp.gov.br/praias/publicacoes-relatorios/)).

112 The collected water samples from the São Domingos Stream represented a poor WQI
113 and a history of poor-quality indices. In contrast, waters obtained from the Pardo River
114 represented the best WQI among all samples and were historically classified as good and
115 excellent. The Tietê River had different WQIs depending on the collection site, with WQI
116 ranging from good to poor. The worst site was TIET-04200, located in São Paulo, the most
117 populous city in Brazil (Table 1; Fig. 1).

118 The first water collections were carried out in March, April, and May 2019. The second
119 collections were carried out in June and July 2019. These two time periods were selected for
120 collection because they corresponded to periods of high and low rainfall in the area,
121 respectively. For each collection, 16 water samples were obtained in triplicate from eight
122 different sites, totaling 48 samples from both sampling periods (Table 1; Fig. 1). Briefly, three
123 liters of water were collected from each site using sterile plastic bottles without preservatives

124 and then transported to the laboratory under refrigeration. Subsequently, all samples were
125 filtered using cellulose membranes 0.22 μm x 25 mm in diameter (Millipore, United States),
126 which were kept at $-20\text{ }^{\circ}\text{C}$ until DNA extraction from the filters.

127

128 *2.2. DNA extraction and untargeted metagenome sequencing*

129

130 Total DNA extraction was performed directly from cellulose membranes using the
131 PowerWater™ DNA Isolation Kit (QIAGEN, United States) according to the manufacturer's
132 recommendations. The DNA concentration and purity were determined using Qubit®
133 Fluorometer (Thermo Fisher Scientific, United States). Triplicate extractions were pooled
134 before sequencing so that 16 samples (8 sites at each of two collection times) were submitted
135 for sequencing. 300 ng of each sample's DNA was sequenced by the untargeted metagenome
136 sequencing method on the Illumina® HiSeq 4000 platform in paired-end mode (2×150 bp)
137 (Illumina Inc., United States) at the Michigan State University Genomics Core Facility facility
138 according to their standard operating protocols. Following the manufacturer's
139 recommendations, the metagenomic libraries were prepared using the TruSeq® Nano DNA
140 Library Preparation Kit with IDT® for Illumina Unique Dual Index adapters (Illumina, United
141 States). Trimmomatic was used to remove adapters and to identify and remove low-quality
142 reads (Phred score value ≤ 33) from raw reads (Bolger et al., 2014). The FASTq files were
143 generated with Illumina Bcl2Fastq v.1.8.4.

144

145 *2.3. Metagenomic data analysis*

146

147 The MicrobeCensus software v1.1.1 (<https://github.com/snayfach/MicrobeCensus>) was
148 used to estimate the average genome size of the microbial communities using metataxonomic

149 data, and the SeqKit platform v.2.2.0 (<https://bioinf.shenwei.me/seqkit/>) was used to generate
150 assembly and read statistics. Microbial diversity was determined using Kraken2 with default
151 parameters (Wood et al., 2019). We used Kraken2 output to calculate the richness and Shannon
152 diversity index with the vegan R package (insert version) (Oksanen et al., 2018). To discover
153 genera correlated with WQI, we used linear models to relate taxon relative abundance to WQI.
154 Model fit (r^2), false discovery rate (FDR; p adjusted value), and slope were calculated for each
155 genus with the function *lm()* in R.

156 The resistome and virulome profiles were determined by ABRicate v.1.0.0
157 (<https://github.com/tseemann/ABRicate>) using ResFinder v.4.0 (Bortolaia et al., 2020), CARD
158 v.3.2.5 (Alcock et al., 2020), and Virulence Factors Database (VFDB) v.2022 (Liu et al., 2022).
159 Genes were detected with a coverage threshold of 60% and a percent identity threshold of 80%.
160 CARD outputs were filtered to select efflux genes related to AMR. The ARGs were categorized
161 at the class level. The efflux genes were grouped by the efflux system pump type. The virulence
162 genes were organized by category due to data sparsity from individual gene counts. Output files
163 from each of the analyses were combined into one data frame, and sample normalization was
164 performed by dividing sample gene abundances by sample genome equivalents calculated in
165 Kraken2. Ecological analyses were performed on the combined and normalized data frame.

166 The Bray-Curtis distance metric was used to calculate dissimilarities between ARGs,
167 efflux genes, and virulence gene abundances detected in river samples. A principal coordinates
168 analysis (PCoA) was performed in the vegan R package v.2.5.6. Only one sample had excellent
169 WQI (PARD-02400-J). This sample was categorized with good WQI samples for downstream
170 analyses (six good, four regular, and six poor WQI samples). The varpart function in vegan was
171 used to calculate variation explained due to river type, WQI, and their combined effect. We
172 tested for significant differences among collection sites, water quality, and the interaction of

173 collection site x water quality using permutational multivariate analysis of variance
174 (PERMANOVA) implemented with the Adonis function in the vegan package.

175 The environmental government agency in São Paulo provided CETESB data regarding
176 river quality metrics. The WQI is a qualitative metric calculated from nine quantitative metrics
177 provided in the CETESB report, including total coliforms, BOD, total nitrogen, total
178 phosphorous, temperature, pH, total solids, total dissolved solids, and dissolved oxygen. We
179 extracted the quantities of each component from CETESB reports, except for total nitrogen due
180 to incomplete data (Supplementary Table S1). Distance-based redundancy analysis using a
181 forward selection approach was performed using the ordi2step function in the vegan package
182 to identify environmental drivers of beta diversity calculated from ARG/efflux/virulence gene
183 abundances.

184

185 *2.4. Data availability*

186 All raw reads were submitted to the NCBI SRA database under the BioProject number
187 PRJNA754506. Bioinformatic and statistic code are available on GitHub
188 (https://github.com/ShadeLab/PAPER_Stehling_riverMetaG_2023).

189

190 **3. Results**

191

192 *3.1. Sequencing data*

193

194 Based on metagenome assembly metrics, the number of paired-end reads and contigs
195 ranged from 29,367,936 to 98,275,048 and 355,111 to 1,428,676 per sequenced sample,
196 respectively. The metagenome with the highest number of genomes ($n=3,801$) was obtained
197 from sample PARD-02500-A, which was collected from the Pardo River. The average genome

198 size ranged from 2,593,349 to 6,510,909, N50 ranged from 662 to 1,074, and the percent GC
199 content ranged from 47.2% to 57.3% (Supplementary Table S2).

200

201 3.2. Microbial diversity

202

203 Microbial diversity analysis revealed the presence of different phyla, classes, orders,
204 and families in all samples. The samples with the highest taxonomic diversity were those
205 collected from the Tietê River, with the highest diversity observed in sample TIET-04200-J,
206 which also had the lowest WQI. Phylum Proteobacteria showed the highest relative abundance
207 in all samples (50.7% to 97.7%), followed by the phyla Actinobacteria and Bacteroidetes, (Fig.
208 2A). At the class level, Actinobacteria, Alphaproteobacteria, Betaproteobacteria, and
209 Gammaproteobacteria were distributed among the samples but had varying proportions (Fig.
210 2B). At the order and family levels, the order Burkholderiales (11.1% to 50.5%) (Fig. 2C) and
211 the family *Comamonadaceae* (11.6% to 33.4%) were the most prevalent groups. Furthermore,
212 there were notable amounts of order Pseudomonadales (Fig. 2C) and Mycobacteriaceae and
213 Oxalobacteraceae families in TIET samples.

214 There was almost no variation in microbial composition between the different collection
215 dates for the sites SDOM-04500, PARD-02400, and PARD-02500. However, the microbial
216 composition of the water samples from the Tietê River had greater collection-to-collection
217 variations from March to July, principally among the samples TIET-02050, TIET-04200, and
218 TIET-03130 (Fig. 2A to 2C). Among these samples, a high diversity of bacterial genera was
219 found. A closer investigation revealed that *Pseudomonas* is prevalent as a potential pathogenic
220 genus in some samples, especially in the poor/regular WQI sample TIET-03130-M (74%).
221 Other pathogenic genera, including *Mycobacterium*, *Burkholderia*, and *Bordetella* were also
222 commonly detected in these river systems but were not in high relative abundance. There were

223 no statistical differences between WQI/sampling month and alpha diversity using
224 PERMANOVA (Fig. 3A and 3B). However, linear models revealed that certain potential
225 opportunistic pathogenic genera [i.e., *Aeromonas* (FDR=0.021, $r^2=0.275$), *Bacteroidetes*
226 (FDR=0.04, $r^2=0.240$), *Pantoea* (FDR=0.026, $r^2=0.256$), and *Parabacteroides* (FDR=0.01,
227 $r^2=0.342$)] had negative correlations with WQI (Fig. 3C).

228

229 3.3. Virulome

230

231 We detected 114 unique virulence genes related to adherence, biofilm, immune
232 modulation, motility, nutritional/metabolic factors, regulation, secretion systems, stress
233 survival, and others. The highest relative abundances of virulence factors were motility (36.7%,
234 $n=572$), adherence (15.6%, $n=218$), and secretion systems (12.8%, $n=179$). In this regard,
235 motility-related genes were mainly detected among the sites from the Tietê River, while
236 regulation-related genes were more prevalent among the sites from the São Domingos Stream.
237 Overall, the Tietê River presented genes related to all virulence factors found. Only genes
238 associated with immune modulation were detected in TIET-02090-J site, while no virulence
239 factor was found at PARD-02500-Jn (good WQI) site (Fig. 4A).

240

241 3.4. Resistome

242

243 Sixty-one unique ARGs were identified. These genes were distributed among different
244 antimicrobial targets, including β -lactams, fluoroquinolones, tetracyclines, aminoglycosides,
245 macrolide-lincosamide-streptogramin (MLS), phenicols, sulfonamides, and trimethoprim.
246 Interestingly, clinically important ARGs were found in all studied aquatic ecosystems (Table
247 2). In general, the highest relative abundances of ARGs were β -lactams (38.6%, $n=51$) and

248 MLS (24.2%, n=32). The abundances of ARGs were higher in samples from the Tietê River,
249 especially in TIET-04200 (poor WQI). Only TIET-04200-J site presented at least one ARG for
250 all antimicrobial classes. Curiously, some sites presented only β -lactams-related ARGs, while
251 PARD-02500-A (regular WQI) did not show ARGs (Fig. 4B).

252 Twenty unique genes associated with nine efflux pumps were detected among the
253 samples. The efflux pumps may confer resistance to multiple antimicrobials, including β -
254 lactams, fluoroquinolones, phenicols, aminoglycosides, tetracyclines, and macrolides. Efflux
255 pumps AcrAB-TolC (39.4%, n=52), MFS-type (18.2%, n=24), MtrCDE (17.4%, n=23), and
256 MexAB-OprM (15.9%, n=21) were the most prevalent. Overall, samples from the Tietê River
257 presented higher relative abundances. Unlike the diversity of virulence genes and ARGs, PARD
258 sites showed only MtrCDE-related genes, and four sites belonging to the Pardo and Tietê rivers
259 did not show any efflux pump-encoding genes (Fig. 4C).

260

261 3.5. PCoA analysis

262

263 A PCoA was constructed to explore dissimilarities between collection sites and
264 virulence/ARG/efflux gene abundances (Fig. 5). Collection type explained 9% of the variation,
265 WQI explained 8% of the variation, and their combined effect explained 12% of the variation.
266 No significance was detected between the abundance of virulence/ARG/efflux genes and
267 collection type/WQI ($0.068 \leq r^2 \leq 0.174$ by Adonis; P values, >0.1). However, the WQI is a
268 qualitative metric calculated from nine quantitative metrics (see Methods). We extracted the
269 quantities of each of these components from environmental reports to perform stepwise model
270 selection using the ordiR2step function. After model selection, it was determined that total
271 solids were a weak predictor of gene abundance patterns ($R^2_{Adj} = 0.078$, P value = 0.04).

272

273 4. Discussion

274

275 In this study, we evaluated three important water sources from the state of São Paulo,
276 Brazil, which presented different WQIs. These parameters were used to test if water quality
277 influences the microbial diversity and the presence of putative pathogenic bacteria and virulome
278 and resistome of these aquatic environments. Regarding the microbial diversity, there was a
279 great predominance of the phylum Proteobacteria in all samples, especially in samples from
280 São Domingos Stream and Tietê River that presented the lowest WQIs. In contrast, samples
281 with higher indices had an increased abundance of Actinobacteria. Despite both phyla have
282 been found in many environments (Wang et al., 2016; Chen et al., 2016; Gilbert et al., 2012),
283 we identified that fluctuations of WQIs impact the abundance of bacterial phyla.

284 In general, analyses of bacterial genera showed that water samples from the Tietê River
285 presented higher relative frequencies for some opportunistic pathogenic bacteria, especially
286 *Pseudomonas*, a ubiquitous and versatile bacterium. Although this genus is not considered
287 commonly abundant within freshwater environments (Newton et al., 2011), it has been isolated
288 from lakes and other water sources and can also be introduced into the microbial community
289 via sewage effluent and stormwater runoff (Batrach et al., 2019; Fresia et al., 2019). In this
290 study, the relative abundance of *Pseudomonas* was significantly higher in most samples
291 collected from the Tietê River in March during the rainy season. This genus is known to use a
292 variety of organic compounds; however, some *Pseudomonas* species, including *P. aeruginosa*,
293 have emerged as opportunistic human pathogens resistant to antimicrobials and/or tolerant to
294 metals (Zhang et al., 2012; Pitondo-Silva et al., 2016; Hwang and Yoon, 2019). Overall, the
295 results indicate the presence of a high diversity of potentially pathogenic bacteria in all samples.

296 Considering the virulome, we identified a high diversity of virulence genes in water
297 samples along the WQI gradient, supporting the notion that aquatic environments serve as

298 reservoirs of virulence factors of pathogenic bacteria (Hamner et al., 2019). Virulence genes
299 were mainly predicted to originate from *P. aeruginosa*, reinforcing the presence of potentially
300 pathogenic *P. aeruginosa* strains in the water samples (Wei et al., 2020). In addition, the other
301 virulence genes found are related to bacterial genera commonly reported to cause infections in
302 humans and animals but also detected in environmental samples (Aziz et al., 2015; Chukwu et
303 al., 2019; Martak et al., 2020).

304 Several studies have been conducted to define the resistome of environmental sources,
305 especially from waters and soils, and ARGs associated with different antimicrobial classes,
306 such as aminoglycosides, tetracyclines, sulfonamides, and β -lactams have been found (Yin et
307 al., 2019; Dang et al., 2020; Qian et al., 2021). These antimicrobials have been commonly used
308 in clinical, aquaculture, and agricultural settings (Qiao et al., 2018). We detected many
309 clinically relevant ARGs, including *bla*_{VEB}, *bla*_{GES}, and *bla*_{OXA-58}, which confer resistance to β -
310 lactams, and *tet(X)*, which is related to tigecycline resistance, both classified as critically
311 important antimicrobials for human and veterinary medicine since they have been used as a last
312 resort to treat infections caused mainly by multidrug-resistant pathogens (Scott et al., 2019;
313 Yang et al., 2020). Although β -lactams are predominantly used to treat human infections,
314 different β -lactamase-encoding genes have been detected in the environment since
315 antimicrobials can reach wastewater and urban sewage mainly through the feces of patients
316 (Auguet et al., 2017; Karkman et al., 2019).

317 In addition, many other ARGs, including *qnrS*, *erm(F)*, *sul*, *cmlA*, and *dfrA*, which are
318 associated with resistance to fluoroquinolones, MFS, sulfonamides, phenicols, and
319 trimethoprim were also found. Notably, tetracyclines and sulfonamides have been extensively
320 used in veterinary medicine as growth promoters and aquaculture production, respectively
321 (Peiris et al., 2017; Song et al., 2017). Consequently, high concentrations of those
322 antimicrobials, as well as their corresponding ARGs, have been extensively detected in different

323 habitats, including rivers (Jia et al., 2018), urban beaches (Furlan et al., 2021), livestock
324 manures (Qian et al., 2018), wastewater treatment plants (Xu et al., 2015; Grabert et al., 2018),
325 and soils (Xiang et al., 2016; Furlan & Stehling, 2021). Interestingly, the highest diversity was
326 found in samples from the Tietê River, mainly in sample TIET-04200-J, which presents the
327 worst WQI and was collected in a dry season. According to some studies, higher concentrations
328 of some antimicrobials are mostly found in dry instead of wet seasons since lack of rainfall may
329 increase antimicrobial concentrations by lack of dilution (Kim et al., 2019; Im et al., 2020).

330 It was expected that virulome/resistome abundances would correlate to river type and/or
331 WQI. However, we did not find this to be the case. Furthermore, total solids were a weak
332 predictor of gene abundance patterns. In this context, total solids can originate from natural
333 sources or human activities, leading to various environmental issues. However, the relationship
334 between total solids and virulome/resistome can be complex and context-dependent. Organic
335 solids can provide surfaces for microorganisms to attach and form biofilms, and others as fecal
336 matter or manure, can contain bacteria carrying especially ARGs that can be released into the
337 water (Holman et al., 2016; Gurmessa et al., 2021).

338

339 **5. Conclusions**

340

341 The metagenomics analysis revealed that *Pseudomonas* was the most abundant detected
342 genus, and despite the wide virulome found, many virulence genes were associated with *P.*
343 *aeruginosa*. Resistome analysis showed genes related to AMR for several antimicrobial classes,
344 highlighting some critically important antimicrobials detected in our study. Some opportunistic
345 pathogenic bacteria had negative correlations with WQI, and no significance was detected
346 between the abundance of virulome/resistome and collection type/WQI, contrasting the
347 hypothesis that only highly polluted environments harbor potentially pathogenic and

348 antimicrobial-resistant bacteria. Therefore, these findings showed that diverse virulome and
349 resistome have been presented in water ecosystems with different WQIs, representing a
350 potential risk to humans, animals, and surrounding environments.

351

352 **Funding**

353 This study was supported by São Paulo Research Foundation (FAPESP; 2018/19539-0 and
354 2021/01655-7) to Eliana Guedes Stehling, and by NSF CAREER 1749544 to Ashley Shade.

355

356 **Declaration of Competing Interest**

357 The authors declare no conflicts of interest.

358

359 **Acknowledgments**

360 Eliana Guedes Stehling, João Pedro Rueda Furlan, and Ralf Lopes thank the FAPESP
361 (2018/01890-3 and 2018/22967-4), the Coordination for the Improvement of Higher Education
362 Personnel (CAPES; 88882.180855/2018-01, 88887.464733/2019-00, and Finance code 001),
363 and the National Council for Scientific and Technological Development (CNPq; 304905/2022-
364 4 and 150712/2022-7) for fellowships. We acknowledge the computational resources from the
365 Michigan State University Institute for Cyber-Enabled Research. Nejc Stopnisek acknowledges
366 support from the Michigan State University Plant Resilience Institute.

367

368 **Ethical Approval**

369 Not required.

370

371 **References**

372

373 Alcock, B.P., Raphenya, A.R., Lau, T.T.Y., Tsang, K.K., Bouchard, M., Edalatmand,
374 A., Huynh, W., Nguyen, A.V., Cheng, A.A., Liu, S., Min, S.Y., Miroshnichenko, A., Tran,
375 H.K., Werfalli, R.E., Nasir, J.A., Oloni, M., Speicher, D.J., Florescu, A., Singh, B., Faltyn, M.,
376 Hernandez-Koutoucheva, A., Sharma, A.N., Bordeleau, E., Pawlowski, A.C., Zubyk, H.L.,
377 Dooley, D., Griffiths, E., Maguire, F., Winsor, G.L., Beiko, R.G., Brinkman, F.S.L., Hsiao,
378 W.W.L., Domselaar, G.V., & McArthur, A.G., 2020. CARD 2020: antibiotic resistance
379 surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.*
380 48(D1):D517-D525. <https://doi.org/10.1093/nar/gkz935>.

381

382 Auguet, O., Pijuan, M., Borrego, C.M., Rodriguez-Mozaz, S., Triadó-Margarit, X.,
383 Giustina, S.V.D., & Gutierrez, O., 2017. Sewers as potential reservoirs of antibiotic resistance.
384 *Sci. Total Environ.* 605-606:1047-1054. <https://doi.org/10.1016/j.scitotenv.2017.06.153>.

385

386 Aziz, R.K., Khalifa, M.M., & Sharaf, R.R., 2015. Contaminated water as a source of
387 *Helicobacter pylori* infection: A review. *J. Adv. Res.* 6:539-547.
388 <https://doi.org/10.1016/j.jare.2013.07.007>.

389

390 Batrich, M., Maskeri, L., Schubert, R., Ho, B., Kohout, M., Abdeljaber, M., Abuhasna,
391 A., Kholoki, M., Psihogios, P., Razzaq, T., Sawhney, S., Siddiqui, S., Xoubi, E., Cooper, A.,
392 Hatzopoulos, T., & Putonti, C., 2019. *Pseudomonas* Diversity Within Urban Freshwaters.
393 *Front. Microbiol.* 10:195. <https://doi.org/10.3389/fmicb.2019.00195>.

394

395 Bolger, A.M., Lohse, M., & Usadel, B., 2014. Trimmomatic: a flexible trimmer for
396 Illumina sequence data. *Bioinformatics.* 30(15):2114-20.
397 <https://doi.org/10.1093/bioinformatics/btu170>.

398

399 Bortolaia, V., Kaas, R.S., Ruppe, E., Roberts, M.C., Schwarz, S., Cattoir, V., Philippon,
400 A., Allesoe, R.L., Rebelo, A.R., Florensa, A.F., Fagelhauer, L., Chakraborty, T., Neumann, B.,
401 Werner, G., Bender, J.K., Stingl, K., Nguyen, M., Coppens, J., Xavier, B.B., Malhotra-Kumar,
402 S., Westh, H., Pinholt, M., Anjum, M.F., Duggett, N.A., Kempf, I., Nykäsenoja, S., Olkkola,
403 S., Wiczorek, K., Amaro, A., Clemente, L., Mossong, J., Losch, S., Ragimbeau, C., Lund, O.,
404 & Aarestrup FM., 2020. ResFinder 4.0 for predictions of phenotypes from genotypes. J.
405 Antimicrob. Chemother. 75(12):3491-3500. <https://doi.org/10.1093/jac/dkaa345>.

406

407 Chen, B., Yuan, K., Chen, X., Yang, Y., Zhang, T., Wang, Y., Luan, T., Zou, S., & Li,
408 X., 2016. Metagenomic Analysis Revealing Antibiotic Resistance Genes (ARGs) and Their
409 Genetic Compartments in the Tibetan Environment. Environ. Sci. Technol. 50:6670-6679.
410 <https://doi.org/10.1021/acs.est.6b00619>.

411

412 Chen, L., Yang, J., Yu, J., Yao, Z., Sun, L., Shen, Y., & Jin, Q., 2005. VFDB: a reference
413 database for bacterial virulence factors. Nucleic Acids Res. 33(Database issue):D325-8.
414 <https://doi.org/10.1093/nar/gki008>.

415

416 Chukwu, M.O., Luther King Abia, A., Ubomba-Jaswa, E., Obi, L., & Dewar, J.B., 2019.
417 Characterization and Phylogenetic Analysis of *Campylobacter* Species Isolated from Paediatric
418 Stool and Water Samples in the Northwest Province, South Africa. Int. J. Environ. Res. Public
419 Health. 16:2205. <https://doi.org/10.3390/ijerph16122205>.

420

421 Dang, C., Xia, Y., Zheng, M., Liu, T., Liu, W., Chen, Q., Ni, J., 2020. Metagenomic
422 insights into the profile of antibiotic resistomes in a large drinking water reservoir. *Environ. Int.*
423 136:105449. <https://doi.org/10.1016/j.envint.2019.105449>.

424

425 Fresia, P., Antelo, V., Salazar, C., Giménez, M., D'Alessandro, B., Afshinnekoo, E.,
426 Mason, C., Gonnet, G.H., & Iraola, G., 2019. Urban metagenomics uncover antibiotic
427 resistance reservoirs in coastal beach and sewage waters. *Microbiome* 7:35.
428 <https://doi.org/10.1186/s40168-019-0648-z>.

429

430 Furlan, J.P.R., & Stehling, E.G., 2021. Multiple sequence types, virulence determinants
431 and antimicrobial resistance genes in multidrug- and colistin-resistant *Escherichia coli* from
432 agricultural and non-agricultural soils. *Env. Pol.* 288:117804.
433 <https://doi.org/10.1016/j.envpol.2021.117804>.

434

435 Furlan, J.P.R., Ramos, M.S., Dos Santos, L.D.R., Gallo, I.F.L., Lopes, R., & Stehling,
436 E.G., 2021. Appearance of *mcr-9*, *bla_{KPC}*, *cfr* and other clinically relevant antimicrobial
437 resistance genes in recreation waters and sands from urban beaches, Brazil. *Mar. Pollut. Bull.*
438 167:112334. <https://doi.org/10.1016/j.marpolbul.2021.112334>.

439

440 Gilbert, J.A., Steele, J.A., Caporaso, J.G., Steinbrück, L., Reeder, J., Temperton, B.,
441 Huse, S., McHardy, A.C., Knight, R., Joint, I., Somerfield, P., Fuhrman, J.A., & Field, D., 2012.
442 Defining seasonal marine microbial community dynamics. *ISME J.* 6:298-308.
443 <https://doi.org/10.1038/ismej.2011.107>.

444

445 Göttig, S., Frank, D., Mungo, E., Nolte, A., Hogardt, M., Besier, S., & Wichelhaus,
446 T.A., 2019. Emergence of ceftazidime/avibactam resistance in KPC-3-producing *Klebsiella*
447 *pneumoniae* in vivo. J. Antimicrob. Chemother. 74:3211-3216.
448 <https://doi.org/10.1093/jac/dkz330>.

449

450 Grabert, R., Boopathy, R., Nathaniel, R., & LaFleur, G., 2018. Effect of tetracycline on
451 ammonia and carbon removal by the facultative bacteria in the anaerobic digester of a sewage
452 treatment plant. Bioresour. Technol. 267:265-270.
453 <https://doi.org/10.1016/j.biortech.2018.07.061>.

454

455 Guo, X., Zhang, R., Liu, J., Sui, Q., & Wei, Y., 2020. Tracking antibiotic resistance
456 gene pollution from different sources using machine learning. Sci. Total Environ. 741:140455.
457 <https://doi.org/10.1016/j.scitotenv.2020.140455>

458

459 Gurmessa, B., Ashworth, A.J., Yang, Y., Savin, M., Moore, P.A.Jr., Ricke, S.C., Corti,
460 G., Pedretti, E.F., & Cocco, S., 2021. Variations in bacterial community structure and
461 antimicrobial resistance gene abundance in cattle manure and poultry litter. Environ. Res.
462 197:111011. <https://doi.org/10.1016/j.envres.2021.111011>

463

464 Hamner, S., Brown, B.L., Hasan, N.A., Franklin, M.J., Doyle, J., Eggers, M.J., Colwell,
465 R.R., & Ford, T.E., 2019. Metagenomic Profiling of Microbial Pathogens in the Little Bighorn
466 River, Montana. Int J Environ Res Public Health. 16:1097.
467 <https://doi.org/10.3390/ijerph16071097>.

468

469 Holman, D.B., Hao, X., Topp, E., Yang, H.E., Alexander, T.W., 2016. Effect of Co-
470 Composting Cattle Manure with Construction and Demolition Waste on the Archaeal,
471 Bacterial, and Fungal Microbiota, and on Antimicrobial Resistance Determinants. PloS One.
472 11(6):e0157539. <https://doi.org/10.1371/journal.pone.0157539>

473

474 Fang, H., Huang, K., Yu, J., Ding, C., Wang, Z., Zhao, C., Yuan, H., Wang, Z., Wang,
475 S., Hu, J., Cui, Y., 2019. Metagenomic analysis of bacterial communities and antibiotic
476 resistance genes in the Eriocheir sinensis freshwater aquaculture environment. Chemosphere.
477 224:202-211. <https://doi.org/10.1016/j.chemosphere.2019.02.068>.

478

479 Hwang, W., & Yoon, S.S., 2019. Virulence Characteristics and an Action Mode of
480 Antibiotic Resistance in Multidrug-Resistant *Pseudomonas aeruginosa*. Sci. Rep. 9:487.
481 <https://doi.org/10.1038/s41598-018-37422-9>.

482

483 Im, J.K., Kim, S.H., Noh, H.R., & Yu, S.J., 2020. Temporal-spatial variation and
484 environmental risk assessment of pharmaceuticals in tributaries of the Han River watershed
485 South Korea. Sci. Total Environ. 741:140486. <https://doi.org/10.1016/j.scitotenv.2020.140486>.

486

487 Jia, J., Guan, Y., Cheng, M., Chen, H., He, J., Wang, S., & Wang, Z., 2018. Occurrence
488 and distribution of antibiotics and antibiotic resistance genes in Ba River, China. Sci. Total
489 Environ. 642:1136-1144. <https://doi.org/10.1016/j.scitotenv.2018.06.149>.

490

491 Jutkina, J., Marathe, N.P., Flach, C. F., Larsson, D.G.J., & Jernberg, C., 2016.
492 Antibiotics and common antibacterial biocides stimulate horizontal transfer of resistance at low
493 concentrations. Sci. Total Environ. 585:64–66. <https://doi.org/10.1016/j.scitotenv.2017.01.136>

494

495 Karkman, A., Pärnänen, K., & Larsson, D.G.J., 2019. Fecal pollution can explain
496 antibiotic resistance gene abundances in anthropogenically impacted environments. *Nat.*
497 *Commun.* 10:80. <https://doi.org/10.1038/s41467-018-07992-3>.

498

499 Kim, B., Ji, K., Kim, C., Kang, H., Lee, S., Kwon, B., Kho, Y., Park, K., Kim, K., &
500 Choi, K., 2019. Pharmaceutical residues in streams near concentrated animal feeding operations
501 of Korea – Occurrences and associated ecological risks. *Sci. Total Environ.* 655:408-413.
502 <https://doi.org/10.1016/j.scitotenv.2018.11.233>.

503

504 Kumari, A., Maurya, N.S., & Tiwari, B., 2020. Hospital wastewater treatment scenario
505 around the globe. *Current Developments in Biotechnology and Bioengineering.* 549-570.
506 <https://doi.org/10.1016/B978-0-12-819722-6.00015-8>.

507

508 Li, D., Liu, C.M., Luo, R., Sadakane, K., & Lam, T.W., 2015. MEGAHIT: an ultra-fast
509 single-node solution for large and complex metagenomics assembly via succinct de Bruijn
510 graph. *Bioinformatics.* 31:1674-1676. <https://doi.org/10.1093/bioinformatics/btv033>.

511

512 Li, B., Yang, Y., Ma, L., Ju, F., Guo, F., Tiedje, J. M., & Zhang, T., 2015. Metagenomic
513 and network analysis reveal wide distribution and co-occurrence of environmental antibiotic
514 resistance genes. *ISME J.* 9:2490–2502. <https://doi.org/10.1038/ismej.2015.59>

515

516 Liu, B., Zheng, D., Zhou, S., Chen, L., & Yang, J., 2022. VFDB 2022: a general
517 classification scheme for bacterial virulence factors. *Nucleic Acids Res.* 50(D1):D912-D917.
518 <https://doi.org/10.1093/nar/gkab1107>

519

520 Loucks, D.P., & van Beek, E., 2017. Water Resources Planning and Management: An
521 Overview. In: Water Resource Systems Planning and Management. Springer Cham.
522 https://doi.org/10.1007/978-3-319-44234-1_1.

523

524 Majumder, A., Gupta, A.K., Ghosal, P.S., & Varma, M., 2021. A review on hospital
525 wastewater treatment: A special emphasis on occurrence and removal of pharmaceutically
526 active compounds, resistant microorganisms, and SARS-CoV-2. J. Environ. Chem. Eng.
527 9:104812. <https://doi.org/10.1016/j.jece.2020.104812>.

528

529 Martak, D., Henriot, C.P., Broussier, M., Couchoud, C., Valot, B., Richard, M.,
530 Couchot, J., Bornette, G., Hocquet, D., & Bertrand, X., 2020. High Prevalence of Human-
531 Associated *Escherichia coli* in Wetlands Located in Eastern France. Front. Microbiol.
532 11:552566. <https://doi.org/10.3389/fmicb.2020.552566>.

533

534 Muniz, P., Venturini, N., Martins, C.C., Munshi, A.B., García-Rodríguez, F., Brugnoli,
535 E., Dauner, A.L.L., Bicego, M.C.B., & García-Alonso, J., 2015. Integrated assessment of
536 contaminants and monitoring of an urbanized temperate harbor (Montevideo, Uruguay): a 12-
537 year comparison. Braz. J. Oceanogr. 63(3).

538

539 Nayfach, S., & Pollard, K.S., 2015. Average genome size estimation improves
540 comparative metagenomics and sheds light on the functional ecology of the human microbiome.
541 Genome Biol. 16:51. <https://doi.org/10.1186/s13059-015-0611-7>.

542

543 Newton, R.J., Jones, S.E., Eiler, A., McMahon, K.D., & Bertilsson, S., 2011. A guide
544 to the natural history of freshwater lake bacteria. *Microbiol. Mol. Biol. Rev.* 75:14-49.
545 <https://doi.org/10.1128/MMBR.00028-10>.

546

547 Nordmann, P., Naas, T., & Poirel, L., 2011. Global spread of carbapenemase-producing
548 *Enterobacteriaceae*. *Emerg. Infect. Dis.* 17:1791-1798.
549 <https://doi.org/10.3201/eid1710.110655>.

550

551 Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D.,
552 Minchin, R.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., &
553 Wagner, H., 2018. Vegan: community ecology package. Rpackage version 2.5–2.
554 <https://CRAN.R-project.org/package=vegan>.

555

556 Ortúzar, M., Esterhuizen, M., Olicón-Hernández, D.R., González-López, J., & Aranda,
557 E., 2022. Pharmaceutical pollution in aquatic environments: A concise review of environmental
558 impacts and bioremediation systems. *Front. Microbiol.* 13:869332.
559 <https://doi.org/10.3389/fmicb.2022.869332>.

560

561 Peiris, C., Gunatilake, S.R., Mlsna, T.E., Mohan, D., & Vithanage, M., 2017. Biochar
562 based removal of antibiotic sulfonamides and tetracyclines in aquatic environments: A critical
563 review. *Bioresour. Technol.* 246:150-159. <https://doi.org/10.1016/j.biortech.2017.07.150>.

564

565 Pitondo-Silva, A., Gonçalves, G.B., & Stehling, E.G., 2016. Heavy metal resistance and
566 virulence profile in *Pseudomonas aeruginosa* isolated from Brazilian soils. *APMIS.* 124:681-
567 688. <https://doi.org/10.1111/apm.12553>.

568

569 Pruden, A., Larsson, D.G.J., Amézquita, A., Collignon, P., Brandt, K.K., Graham, D.
570 W., Lazorchak, J.M., Suzuki, S., Silley, P., & Snape, J. R., 2013. Management options for
571 reducing the release of antibiotics and antibiotic resistance genes to the environment. *Environ.*
572 *Health Perspect.* 121:878–885. <https://doi.org/10.1289/ehp.1206446>

573

574 Qian, X., Gu, J., Sun, W., Wang, X.J., Su, J.Q., & Stedfeld, R. 2018. Diversity,
575 abundance, and persistence of antibiotic resistance genes in various types of animal manure
576 following industrial composting. *J. Hazard Mater.* 344:716-722.
577 <https://doi.org/10.1016/j.jhazmat.2017.11.020>.

578

579 Qian, X., Gunturu, S., Guo, J., Chai, B., Cole, J.R., Gu, J., & Tiedje, J.M., 2021.
580 Metagenomic analysis reveals the shared and distinct features of the soil resistome across
581 tundra, temperate prairie, and tropical ecosystems. *Microbiome.* 9:108.
582 <https://doi.org/10.1186/s40168-021-01047-4>.

583

584 Qiao, M., Ying, G-G., Singer, A.C., & Zhu, Y-G., 2018. Review of antibiotic resistance
585 in China and its environment. *Environ. Int.* 110:160-172. [https://doi.org/10.1016/j.](https://doi.org/10.1016/j.envint.2017.10.016)
586 [envint.2017.10.016](https://doi.org/10.1016/j.envint.2017.10.016).

587

588 Scott, H.M., Acuff, G., Bergeron, G., Bourassa, M.W., Gill, J., Graham, D.W., Kahn,
589 L.H., Morley, P.S., Salois, M.J., Simjee, S., Singer, R.S., Smith, T.C., Storrs, C., & Wittum,
590 T.E., 2019. Critically important antibiotics: criteria and approaches for measuring and reducing
591 their use in food animal agriculture. *Ann. N. Y. Acad. Sci.* 1441(1):8-16.
592 <https://doi.org/10.1111/nyas.14058>

593

594 Song, C., Li, L., Zhang, C., Kamira, B., Qiu, L., Fan, L., Wu, W., Meng, S., Hu, G., &
595 Chen, J., 2017. Occurrence and human dietary assessment of sulfonamide antibiotics in cultured
596 fish around Tai Lake, China. *Environ Sci. Pollut. Res. Int.* 24:17493-17499.
597 <https://doi.org/10.1007/s11356-017-9442-2>.

598

599 Stalder, T., Press, M. O., Sullivan, S., Liachko, I., Top, E.M., & Link, H., 2019. From
600 microbial communities to new biotechnological applications. *Environ. Microbiol.* 21:3–18.
601 <https://doi.org/10.1111/1462-2920.14489>

602

603 Wang, N.F., Zhang, T., Yang, X., Wang, S., Yu, Y., Dong, L.L., Guo, Y.D., Ma, Y.X.,
604 & Zang, J.Y., 2016. Diversity and Composition of Bacterial Community in Soils and Lake
605 Sediments from an Arctic Lake Area. *Front. Microbiol.* 7:1170.
606 <https://doi.org/10.3389/fmicb.2016.01170>.

607

608 Wei, L., Wu, Q., Zhang, J., Guo, W., Gu, Q., Wu, H., Wang, J., Lei, T., Xue, L., Zhang,
609 Y., Wei, X., & Zeng, X., 2020. Prevalence, Virulence, Antimicrobial Resistance, and Molecular
610 Characterization of *Pseudomonas aeruginosa* Isolates From Drinking Water in China. *Front.*
611 *Microbiol.* 11:544653. <https://doi.org/10.3389/fmicb.2020.544653>.

612

613 Winker, M., Schropp, D., Huwe, J., & Ternes, T.A., 2020. Fate and spatial distribution
614 of antibiotics, antibiotic resistant bacteria and antibiotic resistance genes in manure-amended
615 soils: A critical review of studies over the last 20 years. *Sci. Total Environ.* 744:140822.
616 <https://doi.org/10.1016/j.scitotenv.2020.140822>

617

618 Wood, D.E., Lu, J., & Langmead, B., 2019. Improved metagenomic analysis with
619 Kraken 2. *Genome Biol.* 20(1):257. <https://doi.org/10.1186/s13059-019-1891-0>.

620

621 Xiang, L., Wu, X.L., Jiang, Y.N., Yan, Q.Y., Li, Y.W., Huang, X.P., Cai, Q.Y., & Mo,
622 C.H., 2016. Occurrence and risk assessment of tetracycline antibiotics in soil from organic
623 vegetable farms in a subtropical city, south China. *Environ. Sci. Pollut. Res. Int.* 23:13984-
624 13995. <https://doi.org/10.1007/s11356-016-6493-8>.

625

626 Xu, J., Xu, Y., Wang, H., Guo, C., Qiu, H., He, Y., Zhang, Y., Li, X., & Meng, W.,
627 2015. Occurrence of antibiotics and antibiotic resistance genes in a sewage treatment plant and
628 its effluent-receiving river. *Chemosphere.* 119:1379-1385.
629 <https://doi.org/10.1016/j.chemosphere.2014.02.040>.

630

631 Yang, Y., Li, B., Zou, S., Fang, H.H., & Zhang, T., 2014. Fate of antibiotic resistance
632 genes in sewage treatment plant revealed by metagenomic approach. *Water Res.* 62:97–106.
633 <https://doi.org/10.1016/j.watres.2014.05.019>

634

635 Yang, Q., Pogue, J.M., Li, Z., Nation, R.L., Kaye, K.S., Li, J., 2020. Agents of Last
636 Resort: An Update on Polymyxin Resistance. *Infect. Dis. Clin. North Am.* 34(4):723-750.
637 <https://doi.org/10.1016/j.idc.2020.08.003>.

638

639 Yin, X., Deng, Y., Ma, L., Wang, Y., Chan, L.Y.L., & Zhang, T., 2019. Exploration of
640 the antibiotic resistome in a wastewater treatment plant by a nine-year longitudinal
641 metagenomic study. *Environ. Int.* 133(Pt B):105270.
642 <https://doi.org/10.1016/j.envint.2019.105270>.

643

644 Zhang, Q-Q., Ying, G-G., Pan, C-G., Liu, Y-S., & Zhao, J-L., 2015. Comprehensive
645 Evaluation of antibiotics emission and fate in the river basins of China: source analysis,
646 multimedia modeling, and linkage to bacterial resistance. *Environ. Sci. Technol.* 49:6772-6782.
647 <https://doi.org/10.1021/acs.est.5b00729>.

648

649 Zhang, Y., Marrs, C. F., Simon, C., & Xi, C. (2009). Wastewater treatment contributes
650 to selective increase of antibiotic resistance among *Acinetobacter* spp. *Sci. Total Environ.*
651 407:3702–3706. <https://doi.org/10.1016/j.scitotenv.2009.02.040>

652

653 Zhang, Y.X., Wang, J., Chai, T.Y., Zhang, Q., Liu, J.G., Li, X., Bai, Z.Q., & Su, Z.J.,
654 2012. [Mechanism of heavy-metal tolerance in *Pseudomonas aeruginosa* ZGKD2]. *Huan Jing*
655 *Ke Xue.* 33:3613-3619.

656 **Captions of tables and figures**

657

658 **Table 1.** This study sampled the field data, sample collection, and WQIs of the stream
659 environments.

660

661 **Table 2.** The most clinically important ARGs that were found for each antimicrobial class
662 among the stream samples.

663

664 **Fig. 1.** A map showing the collection sites of the aquatic environments. Collection sites: Pardo
665 River, PARD; São Domingos Stream, SDOM; Tietê River, TIET.

666

667 **Fig. 2. Fig. 2.** Relative abundances of the microbial taxonomic groups found in each water
668 sample, with reference to each WQI (colored circles on the bottom). (A) Phylum; (B) Class;
669 (C) Order. Collection sites: Pardo River, PARD; São Domingos Stream, SDOM; Tietê River,
670 TIET.

671

672 **Fig 3.** Impact of environmental factors on the alpha diversity and abundance of potential
673 opportunistic pathogens. Alpha diversity, represented by richness (A) and the Shannon diversity
674 index (B), is plotted against the time of sample collection [early (March, April, May) and late
675 (June, July)]. Points from the same sampling locations are connected by a line and colored by
676 the WQI. We used linear models to find genera significantly correlated with the WQI. Four of
677 61 genera negatively correlated with the WQI belonged to potential opportunistic pathogenic
678 genera (C). Collection sites: Pardo River, PARD; São Domingos Stream, SDOM; Tietê River,
679 TIET.

680

681 **Fig. 4.** Relative frequency of genes encoding virulence factors (A), antimicrobial resistance (B),
682 and efflux pumps (C) found in each sample. Collection sites: Pardo River, PARD; São
683 Domingos Stream, SDOM; Tietê River, TIET.

684

685 **Fig. 5.** A combined PCoA to explore the dissimilarities between the collection sites and
686 abundances of virulence genes, ARGs, and efflux pump-encoding genes. Collection sites: Pardo
687 River, PARD; São Domingos Stream, SDOM; Tietê River, TIET.

Table 1. Field data, sample collection, and WQIs of the stream environments sampled as part of this study

Source	Collection site	Sample	Collection date	GPS data	City ¹	WQI (value) ²
Pardo River	PARD-02400	PARD-02400-A	4/23/19	21 17 32S, 47 28 52W	Santa Cruz da Esperança	Good (73)
		PARD-02400-Jn	6/6/19			Excellent (80)
	PARD-02500	PARD-02500-A	4/16/19	21 06 00S, 47 45 44W	Ribeirão Preto	Regular (51)
		PARD-02500-Jn	6/26/19			Good (65)
São Domingos Stream	SDOM-04500	SDOM-04500-M	5/16/19	21 03 02S, 49 03 49W	Catiguá	Poor (36)
		SDOM-04500-J	7/18/19			Poor (27)
Tietê River	TIET-02050	TIET-02050-M	3/7/19	23 33 54S, 46 00 57W	Biritiba Mirim	Good (55)
		TIET-02050-J	7/26/19			Good (76)
	TIET-02090	TIET-02090-M	3/7/19	23 32 55S, 46 08 09W	Mogi das Cruzes	Regular (51)
		TIET-02090-J	7/26/19			Good (67)
	TIET-03130	TIET-03130-M	3/7/19	23 28 19S, 46 20 50W	Itaquaquecetuba	Regular (37)
		TIET-03130-J	7/26/19			Poor (35)
	TIET-04200	TIET-04200-M	3/7/19	23 31 11S, 46 44 47W	São Paulo	Poor (36)
		TIET-04200-J	7/26/19			Poor (21)
	TIET-02400	TIET-02400-M	3/7/19	23 05 12S, 47 40 41W	Tietê	Poor (32)
		TIET-02400-J	7/26/19			Regular (38)

¹ All cities belonging to the state of São Paulo.

² WQI values (Poor: 20 to 36; Regular: 37 to 51; Good: 52 to 79; Excellent: 80 to 100) according to public reports of CETESB available at (<https://cetesb.sp.gov.br/praias/publicacoes-relatorios/>)

Table 2. The most clinically important ARGs found for each antimicrobial class among the stream samples

Gene	Antimicrobial target¹	Source	Species found²
<i>bla_{GES}</i>	β-lactams	Tietê River	Enterobacteria, <i>A. baumannii</i> , <i>P. aeruginosa</i>
<i>bla_{VEB}</i>		Tietê River	<i>E. coli</i> , <i>K. pneumoniae</i> , <i>A. baumannii</i> , <i>P. aeruginosa</i>
<i>bla_{OXA-58}</i>		Tietê River	<i>K. pneumoniae</i> , <i>P. mirabilis</i> , <i>Acinetobacter</i> sp.
<i>qnrS</i>	Fluoroquinolones	Tietê River	Enterobacteria, <i>A. baumannii</i>
<i>aph(3'')-Ib</i>	Aminoglycosides	Tietê River	Enterobacteria, NFGNB
<i>tet(A)</i>	Tetracyclines	Tietê River	Enterobacteria, NFGNB
<i>tet(B)</i>		Tietê River	Enterobacteria, NFGNB
<i>tet(X)</i>		Tietê River	<i>E. coli</i> , <i>Bacteroides</i> sp., <i>Myroides</i> sp., <i>Phocaeicola</i> sp.
<i>erm(F)</i>	MLS	São Domingos Stream, Tietê River	<i>Bacteroides</i> sp., <i>Myroides</i> sp., <i>Phocaeicola</i> sp.
<i>sul1</i>	Sulfonamides	São Domingos Stream, Pardo River, Tietê River	Enterobacteria, NFGNB
<i>sul2</i>		São Domingos Stream, Tietê River	Enterobacteria, NFGNB
<i>cmlA</i>	Phenicols	Tietê River	Enterobacteria, <i>P. aeruginosa</i>
<i>dfrA</i>	Trimethoprim	Tietê River	Enterobacteria, <i>A. baumannii</i> , <i>P. aeruginosa</i>

¹ Macrolide-lincosamide-streptogramin, MLS. ² Description according to resistomes with perfect matches and/or resistomes with sequence variants available in the CARD database; Non-fermenting Gram-negative bacilli, NFGNB.

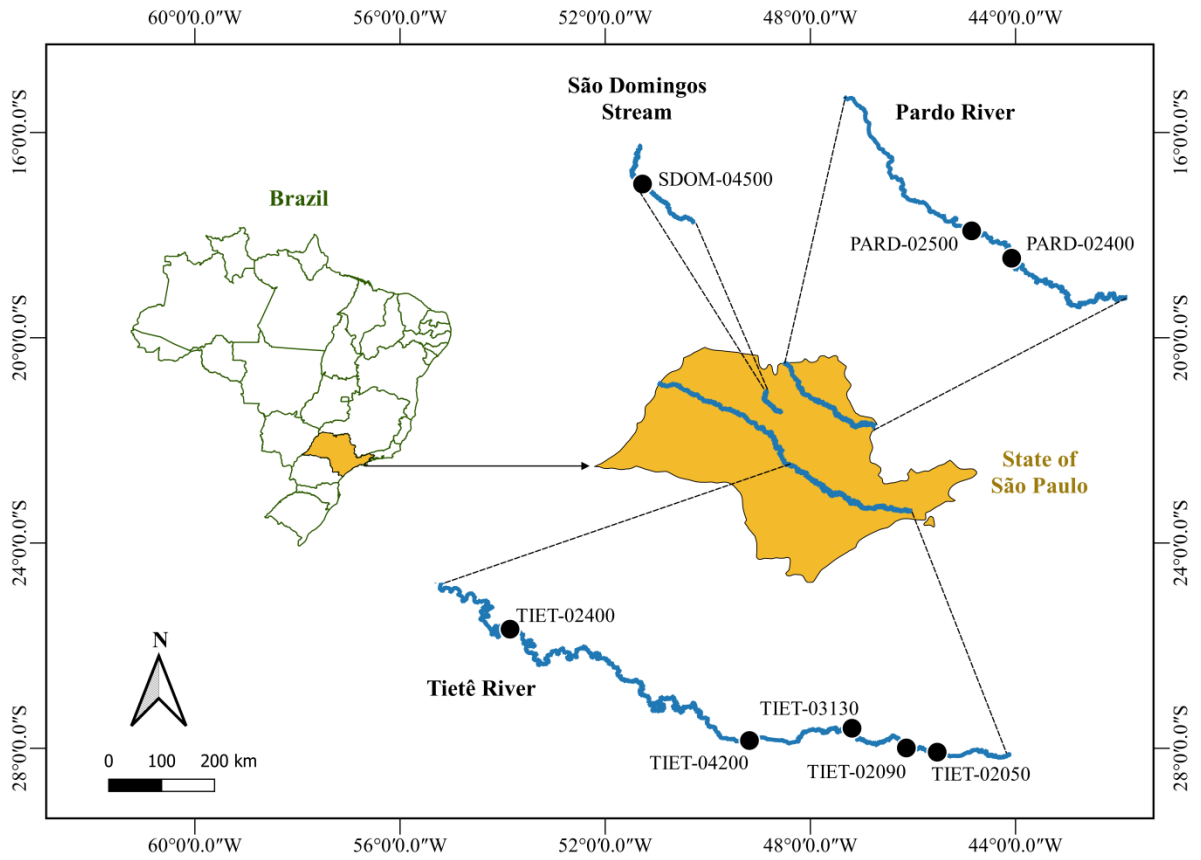


Fig. 1. Map showing the collection sites of the aquatic environments. Collection sites: Pardo River, PARD; São Domingos Stream, SDOM; Tietê River, TIET.

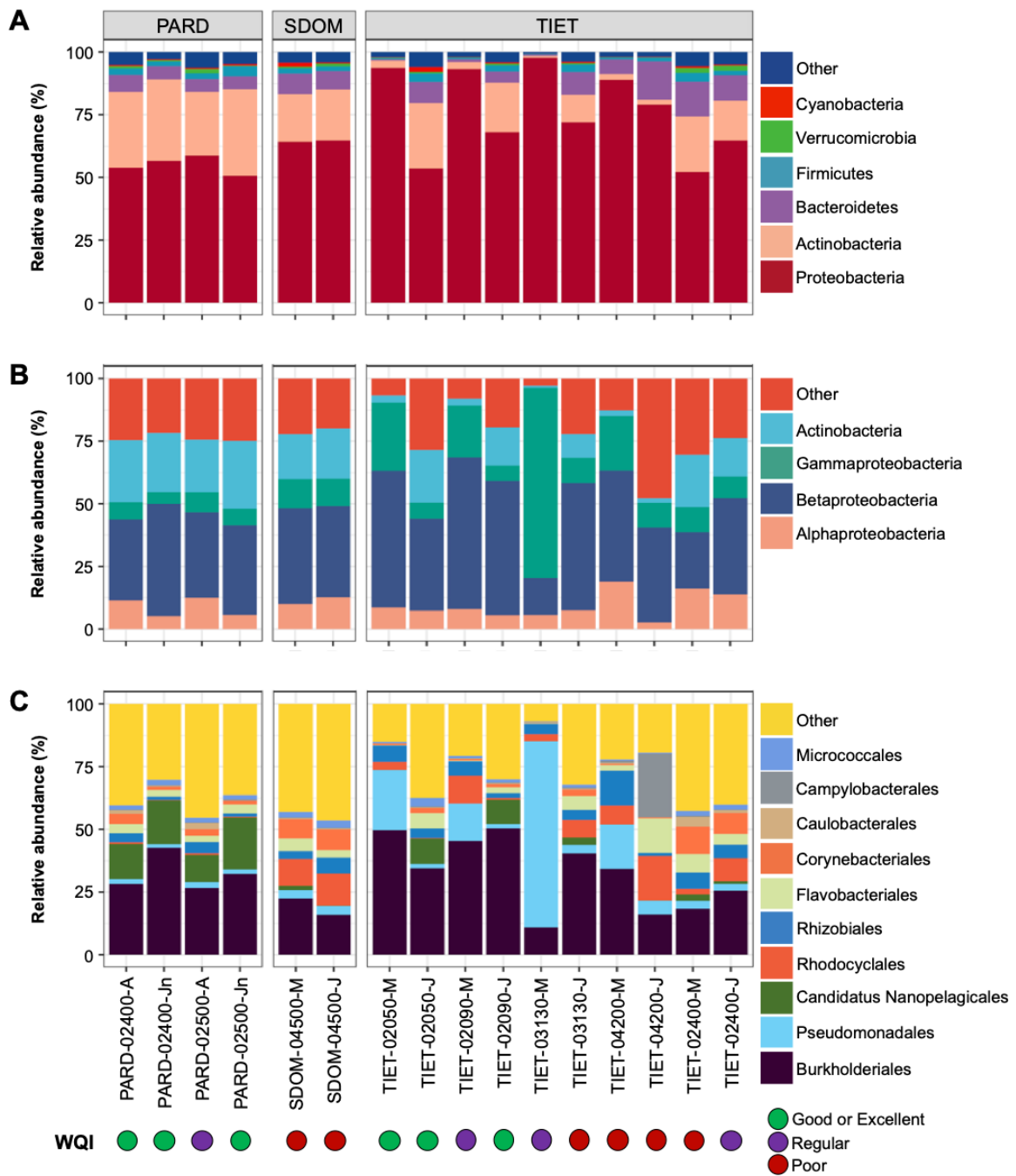


Fig. 2. Relative abundances of the microbial taxonomic groups found in each water sample, with reference to each WQI (colored circles on the bottom). (A) Phylum; (B) Class; (C) Order. Collection sites: Pardo River, PARD; São Domingos Stream, SDOM; Tietê River, TIET.

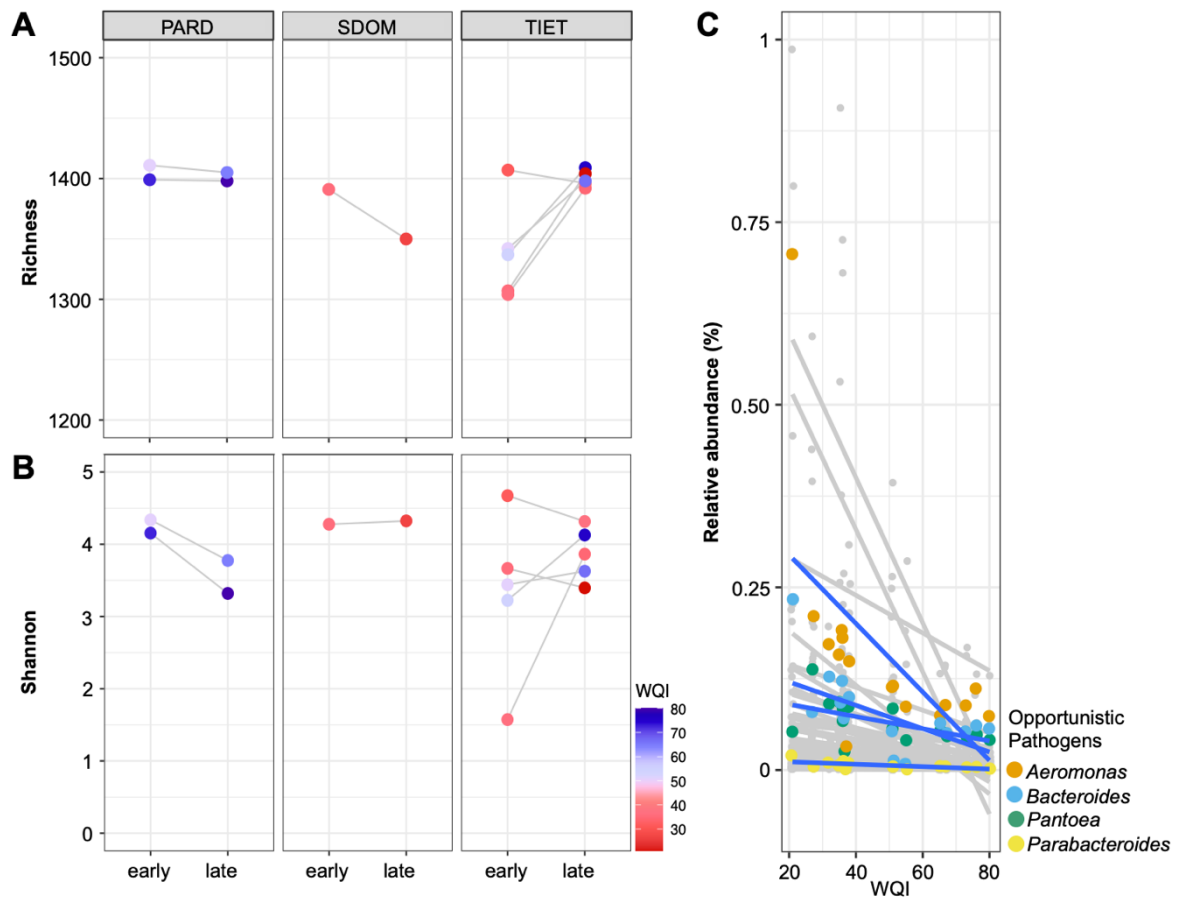


Fig 3. Impact of environmental factors on the alpha diversity and abundance of potential opportunistic pathogens. Alpha diversity, represented by richness (A) and the Shannon diversity index (B), is plotted against the time of sample collection [early (March, April, May) and late (June, July)]. Points from the same sampling locations are connected by a line and colored by the WQI. We used linear models to find genera significantly correlated with the WQI. Four of 61 genera negatively correlated with the WQI belonged to potential opportunistic pathogenic genera (C). Collection sites: Pardo River, PARD; São Domingos Stream, SDOM; Tietê River, TIET.

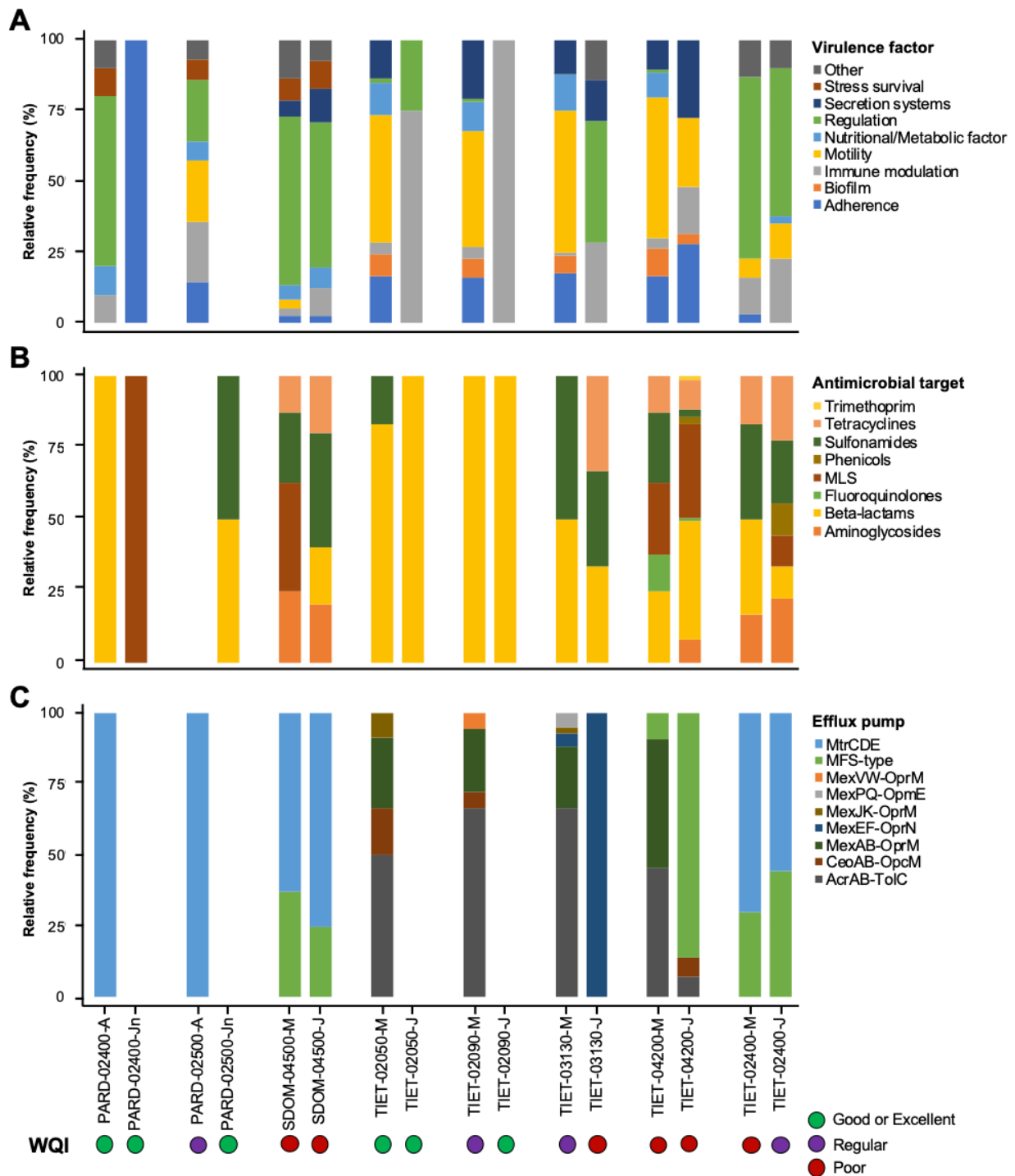


Fig. 4. Relative frequency of genes encoding virulence factors (A), antimicrobial resistance (B), and efflux pumps (C) found in each sample. Collection sites: Pardo River, PARD; São Domingos Stream, SDOM; Tietê River, TIET.

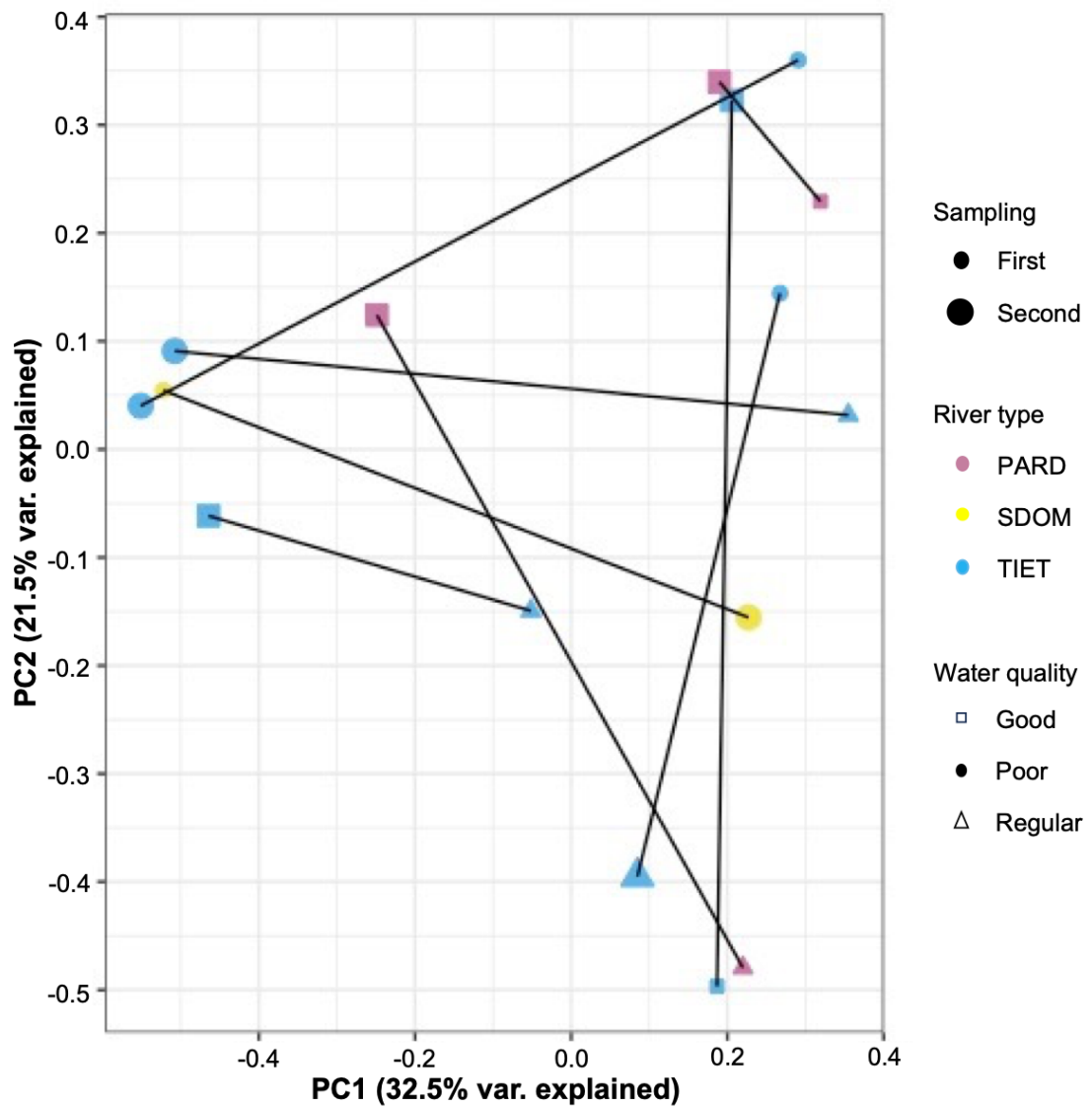


Fig. 5. A combined PCoA to explore the dissimilarities between the collection sites and abundances of virulence genes, ARGs, and efflux pump-encoding genes. Collection sites: Pardo River, PARD; São Domingos Stream, SDOM; Tietê River, TIET.