This document is the accepted manuscript version of the following article:
Lee, J., Ju, F., Maile-Moskowitz, A., Beck, K., Maccagnan, A., McArdell, C. S., … Bürgmann, H. (2021). Unraveling the riverine antibiotic resistome: the downstream fate of anthropogenic inputs. Water Research. https://doi.org/10.1016/j.watres.2021.117050

Unraveling the riverine antibiotic resistome: the downstream fate of anthropogenic inputs

Jangwoo Lee[1,2], Feng Ju[3,4], Ayella Maile-Moskowitz[5], Karin Beck[1], Andreas Maccagnan[1], Christa S. McArdell[1], Marco Dal Molin[1,6], Fabrizio Fenicia[1], Peter Vikesland[5], Amy Pruden[5], Christian Stamm[1], Helmut Bürgmann[1,*]

[1]Eawag, Swiss Federal Institute of Aquatic Science and Technology, CH-6047 Kastanienbaum or CH-8600 Dübendorf, Switzerland

[2]Department of Environmental Systems Science, ETH Zurich, Swiss Federal Institute of Technology, Zurich, Switzerland

[3]Key Laboratory of Coastal Environment and Resources of Zhejiang Province, School of Engineering, Westlake University, Hangzhou, China

[4]Institute of Advanced Technology, Westlake Institute for Advanced Study, Hangzhou, China

[5]Department of Civil and Environmental Engineering, Virginia Tech, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, United States of America

[6]The Centre of Hydrogeology and Geothermics (CHYN), University of Neuchâtel, 2000 Neuchâtel, Switzerland

* Corresponding author: Helmut Bürgmann, Eawag, Seestrasse 79, 6047 Kastanienbaum, Switzerland. E-mail: Helmut.Buergmann@eawag.ch

Highlights

- Initial rapid decrease of wastewater-borne resistance levels is primarily explained by dilution
- Additional source/sink effects become apparent over longer downstream distances
- Correct evaluation of resistance determinant fate requires mass-transport analysis
- *bacA* was prevalent in less-disturbed waters proving its high natural prevalence

Abstract

River networks are one of the main routes by which the public could be exposed to environmental sources of antibiotic resistance, that may be introduced e.g. via treated wastewater. In this study, we applied a comprehensive integrated analysis encompassing mass-flow concepts, chemistry, bacterial plate counts, resistance gene quantification and shotgun metagenomics to track the fate of the resistome (collective antibiotic resistance genes (ARGs) in a microbial community) of treated wastewater in two Swiss rivers at the kilometer scale. The levels of certain ARGs and the class 1 integron integrase gene (*intI1*) commonly associated with anthropogenic sources of ARGs decreased quickly over short distances (2-2.5 km) downstream of wastewater discharge points. Mass-flow analysis based on conservative tracers suggested this decrease was attributable mainly to dilution but ARG loadings frequently also decreased (e.g., 55.0-98.5 % for *ermB* and *tetW*) over the longest studied distances (6.8 and 13.7 km downstream). Metagenomic analysis confirmed that ARG of wastewater-origin did not persist in rivers after 5 ~ 6.8 km downstream distance. *sul1* and *intI1* levels and loadings were more variable and even increased sharply at 5 ~ 6.8 km downstream distance on one occasion. While input from agriculture and in-situ positive selection pressure for organisms carrying ARGs cannot be excluded, in-system growth of biomass is a more probable explanation. The potential for direct human exposure to the resistome of wastewater-origin thus appeared to typically abate rapidly in the studied rivers. However, the riverine aquatic resistome was also dynamic, as evidenced by the increase of certain gene markers downstream, without obvious sources of anthropogenic contamination. This study provides new insight into drivers of riverine resistomes and pinpoints key monitoring targets indicative of where human sources and exposures are likely to be most acute.

Keywords: Antimicrobial resistance; Wastewater; River system; Metagenomics; Transport; Degradation

1. Introduction

Antibiotic resistance is increasingly recognized by international and governmental entities as a growing global public health threat. According to a 2014 report by the Wellcome Trust and the British government, more than 50,000 cases of antibiotic resistant infections occur annually in Europe and the United States and many hundreds of thousands of people die due to infections with resistant bacteria in other regions of the globe (O'Neill, 2014). In the EU and European Economic Area, the annual attributable deaths by infection with antibiotic resistant pathogens have increased significantly between 2007 and 2015, for instance, from 11,000 to 27,000 (Cassini et al., 2019).

Aquatic environments play a potentially important role as routes of dissemination of resistance; environmental niches at the landscape scale are connected to uses including drinking water supply, irrigation, and recreation. Research in this area has greatly intensified over the last decade (Bürgmann et al., 2018; Rizzo et al., 2013; Zhang et al., 2009) and an increasing number of studies have investigated anthropogenic impacts on receiving rivers. Among the earliest investigations were the studies on the Poudre River in Colorado, United States, which proposed quantitative polymerase chain reaction (qPCR)-based quantification of various antibiotic resistance genes (ARGs), along with phylogenetic analysis (e.g., *tetW*), as a framework for tracking anthropogenic inputs. Anthropogenic input of ARGs to the receiving river was well-apparent using this approach (Storteboom et al., 2010). More recently, the advent of shotgun metagenomic sequencing has greatly advanced the resolution in the ability to characterize large-scale impacts of anthropogenic ARG inputs, as was observed in the Han river catchment in Korea (Lee et al., 2020). The authors noted a strong association of fecal contamination as evidence of anthropogenic activities shaping the composition of the downstream antibiotic resistome (collective ARGs in a microbial community). In Switzerland, a study on rivers and lakes identified the occurrence of extended spectrum β lactamase- and carbapenemase-producing *Enterobacteriaceae*, which presumably originated from anthropogenic activities (Zurfluh et al., 2013). Another recent study revealed that stream microbiota are significantly altered by the input of treated wastewater in natural streams (Mansfeldt et al., 2020).

ARGs and resistant bacteria (ARB) can persist or proliferate in environmental systems by various mechanisms. Horizontal gene transfer may occur, potentially resulting in new combinations of ARGs or the transfer of resistance to environmentally-adapted bacteria that could in turn change the role of the environment as reservoirs of resistance for clinically-relevant bacteria. Furthermore, the possibility of resistance selection under sub-inhibitory concentrations of antibiotics has been reported (Andersson and Hughes, 2012). Recently, first attempts have been made to estimate predicted no effect concentrations (PNECs) for resistance selection in environmental settings (Bengtsson-Palme and Larsson, 2016). However, current PNECs are an estimate extrapolated from data on isolated bacteria, and could vary substantially under in-situ environmental conditions and with environmental bacteria.

The above examples make clear that treated wastewater discharges have a significant impact on the abundance and types of ARB and ARGs in receiving rivers. Thus, it is crucial that we gain a better understanding of the downstream fate of the anthropogenic antibiotic resistome in receiving rivers. In this sense, few previous studies have attempted to investigate the downstream behavior of various indicators of resistance (e.g., ARBs, ARGs, mobile genetic elements commonly associated with ARGs), and no clear picture of such behavior has as yet emerged. For instance, a study performed in two wastewater treatment plants (WWTPs) and their receiving river in China reported that the levels of wastewater-origin ARGs (*tetC*, *sul1*) and the class 1 integron integrase gene (*intI1*), decreased significantly 1.2 ~ 2.5 km downstream of the wastewater discharge point (Li et al., 2016). On the other hand, a study performed in a Dutch stream showed that the downstream levels of *sul1*, *sul2*, *ermB*, *tetW*, and *intI1* persisted, or even increased for certain genes over a 20 km downstream distance (Sabri et al., 2018). Mass-flow analyses of ARB and ARG are missing. These contradictory results regarding the downstream behavior of resistance determinants could be in principle attributable to various factors – different geo-hydrological conditions, potential inputs from non-point (e.g. agricultural) sources, and the possible existence of biological drivers (i.e., horizontal and/or vertical gene transfer). An improved understanding of the fate of the wastewater-origin antibiotic resistomes and underlying causes would therefore require an integrated approach across multiple disciplines.

The purpose of this study was to track wastewater-origin antibiotic resistomes and identify the key mechanisms governing their fate in two of the most substantially wastewater-impacted rivers in Switzerland. It was hypothesized, that short-distance (up to 1~2 km from wastewater discharge point) behavior of wastewater-origin resistance determinant concentrations would be governed mostly by hydrological effects such as mixing and dilution. Thus, we used conservative chemical tracers to determine dilution effects and further investigated the contribution of dilution on downstream dynamics of resistance determinants. On the other hand, it was expected that over longer distances (more than 1~2 km; up to 13.7 km to the next downstream WWTP) fate of ARGs and ARB would depend also on additional source/sink mechanisms, such as inputs from biological drivers (e.g., death or growth of wastewater-origin ARB, in-situ resistance (co)selection by antibiotics or metals, and horizontal and/or vertical gene transfer), and/or non-point sources (e.g., agricultural runoff) which are expected to diffuse into the system continuously from a large catchment area. Therefore, the potential effects of biological drivers over long downstream distances were investigated after accounting for hydrological effects. To provide a comprehensive assessment of indicators of antibiotic resistance in the environment, we combined various approaches: cultivation of heterotrophic bacteria on media containing antibiotics, quantification of key indicators of anthropogenic sources of antibiotic resistance by qPCR, and broad profiling of the resistome in select samples using shotgun metagenomic sequencing. Our study contributes to a systematic, interdisciplinary understanding of the mechanisms driving the fate of the wastewater antibiotic resistome in anthropogenically-impacted rivers.

2. Material and methods

2.1. Site description and field work

A list of WWTP effluent-receiving Swiss rivers without known upstream point-source inputs (e.g., other WWTPs) was obtained from a database provided by Eawag, the Swiss Federal Institute of Aquatic Science and Technology (retrieved at 2018) (Eawag, 2014). Two sites were selected according to the following criteria: 1) Greatest proportion of effluent discharge to river discharge, 2) Least number of side streams (for minimum dilution effect from side streams), and 3) Longest distance until the receiving river reaches another downstream WWTP. The selected sites were the River Suze

in Villeret (VIL) in canton Bern, and the River Murg in Münchwilen (MUE) in canton Thurgau (Maps with all sampling points, see supplementary Fig. S1 and S2). At the sampled sections, both are shallow (generally <30 cm depth under low flow conditions, maximum depth 1 m) rivers of Strahler order number 3 and 6, and a mean annual runoff of 2.03 and 1.61 m^3 /s, respectively. The river beds are mostly gravel. To avoid elevated flow conditions we sampled only under dry weather conditions at the time of sampling and during at least the previous 36 hours.

At VIL, we studied a 23.7 km stretch of the Suze that we sampled from 10 km upstream (US5) of the effluent (EF) discharge point of WWTP Villeret and at 8 downstream sites located from 0.5 km (D1) to 13.7 km flow distance downstream (D8) before the Suze reaches the discharge point of another WWTP. Four sampling campaigns were performed in 2018 on July 09 (VI L1), July 19 (VIL2), July 30 (VIL3), and November 05 (VIL4). Different combinations of locations were sampled in each sampling campaign as described in detail in the SI (pp. 4-5). Daily discharge measures from two gauging stations, one near US, and the other near D8 were obtained and are given in Dataset S1.

At MUE, we studied a 7.0 km stretch of the Murg that we sampled from 0.2 km upstream (US) of WWTP Münchwilen and at 8 downstream sites located from 0.5 (D1) to 6.8 km flow distance downstream (D8 before the Murg reaches the discharge point of another WWTP. Three sampling campaigns were performed in 2018 on July 26 (MUE1), August 03 (MUE2), and August 06 (MUE3). Discharge data was obtained from a gauging station near D4 (Dataset S1).

Samplings were performed referring to other projects performed in Swiss rivers and WWTPs (Mansfeldt et al., 2020; Ju et al., 2019). At each river sampling location grab samples (5L in sterilized water containers) were obtained by combining water from just below the surface at three points along a river transect: in the middle and roughly equidistant from the banks to each side of the middle point. EF samples were obtained from the final effluent stream of the WWTPs prior to discharge. Temperature (°C), conductivity (µS/cm), pH, and dissolved oxygen (DO, mg/L), were measured on site in an aliquot of the sample using a portable multi-parameter probe (Multi 3630 IDS, WTW, Germany) at the time of sampling. To make sure EF was fully mixed with receiving water at D1, the conductivity values across cross-section were measured, and no significant deviation was observed (<

0.5%). All samples were cooled at 4 °C in the dark while transported to our laboratory on the same day. Samples were processed on the same and next day within 36 hours. For organic micropollutant analysis, water samples were obtained separately, and stored in pre-combusted glass bottles on site, cooled at 4 °C during transportation, and frozen at -20 °C in the dark upon arrival at the laboratory until analyzed. Sediment samples were obtained from 5 select locations (US, D1, D2, D5, and D8) for select campaigns (VIL1~3 and MUE1~3), and frozen at -20 °C upon arrival at our laboratory.

To better constrain flow velocities in the rivers, salt tracer experiment using NaCl and flow-velocity measurement were performed in separate sampling campaigns in August 2019 (August 23, 2019 for VIL, and August 27, 2019 for MUE) under comparable flow conditions. The results were summarized in Dataset S4 (e.g., flow-velocity and hydraulic residence time).

Further details on sites, field sampling procedures and hydrological experiments are given in the Method section of the SI (pp. 4-5). The exact sampling locations (GPS-coordinates) are given in Table S1.

2.2. Heterotrophic plate count of antibiotic resistant bacteria (ARB)

Levels (colony forming units (CFUs) per mL) of ARB cultivable on R2A agar plates were determined in the presence of two combinations of antibiotics: 1) clarithromycin (4.0 mg/L) and tetracycline (16.0 mg/L) (CLR/TET), and 2) sulfamethoxazole (76.0 mg/L), trimethoprim (4.0 mg/L), and tetracycline (16.0 mg/L) (SMX/TMP/TET) referring to the resistance breakpoints for Enterobacteriaceae suggested by Clinical and Laboratory Standards Institute (Cockerill et al., 2013) and also one of our previous publications (Czekalski et al., 2012). The detailed protocol is available in the SI (pp. 6-7).

2.3. DNA extraction and quantitative PCR

Two aliquots of each water sample were filtered through two 0.2 µm pore size membrane filters, using 0.5 L for EF and 1.0 L for river water samples. Replicate filters were then processed separately. DNA was extracted from the filters using DNeasy PowerWater Kit (Qiagen, Germany) following the manufacturer's instruction. For sediment samples, DNA was extracted from about 20 g of wet

sediment using the DNeasy PowerMax Soil Kit (Qiagen, Germany). Extraction blanks confirmed absence of DNA contamination (see SI, p.8) . The concentration and qualities of extracted DNA were measured using a NanoDrop One spectrophotometer (Thermo Fisher Scientific, USA) (Dataset S2).

Presence and abundance of key indicator genes for anthropogenic ARG inputs (*sul1*, *tetW*, *ermB*, bla_{CTX} and integron integrase class 1 gene) (Berendonk et al., 2015; Gillings et al., 2015;Ju et al., 2019), were determined by qPCR as described previously (Czekalski et al., 2012; Czekalski et al., 2014). The detailed qPCR protocols are reported in the SI. Absence of contamination from filtration and extraction procedures was confirmed using an experiment control by qPCR analysis as shown in the SI.

2.4. Metagenome and 16S rRNA amplicon sequencing analysis

Shotgun metagenomics and 16S rRNA gene amplicon sequencing analysis were performed using Illumina platforms for samples from three selected sampling campaigns. Samples were selected for sequencing according to following rationale: For VIL the samples were selected only from campaign VIL1 (i.e., 6 samples: US, EF, D1, D3, D5, D8) as the samples from other campaigns (VIL2-3) showed similar patterns of resistance determinant profiles downstream. For MUE, 6 samples from MUE2 and MUE3 campaigns (i.e., US, EF, D1, D3, D5, D8 from each sampling) were selected as the far downstream behaviors of certain ARG (e.g., *sul1*) and *intI1* were significantly different from each other in those campaigns. DNA extracts from replicated filters were pooled. All library construction and sequencing was performed by Novogene (Hong Kong).

A detailed description of the bioinformatics workflow is given in the SI. Briefly, metagenomic data were analyzed as follows: 1) After quality controls of metagenome reads, de-novo assembly was performed using MEGAHIT v1.1.3 (Li et al., 2015), 2). Open reading frames (ORFs) were predicted from the assembled contigs using Prodigal v2.6.3 (Hyatt et al., 2010), and annotated to ARGs using the Structured Antibiotic Resistance Genes (SARG) v2.0 database (Yin et al., 2018), 3). After read mapping to contigs and ORFs using Bowtie2 (Langmead and Salzberg, 2012) and Samtools (Li et al., 2009), the coverage information for contigs and ORFs was calculated according to Albertsen et al. (2013). 4) Using the coverage information, abundance metrics were calculated as described in Table

S3. 5) Further downstream analyses were performed, such as contig-based taxonomy assignment using Kaiju v1.7.2 (Menzel et al., 2016) Kraken2 (Wood et al., 2019) and BLASTN, and detailed annotation and visualization of ARG-containing contigs.

To analyze 16S rRNA gene amplicon sequencing data, we used the DADA2 pipeline (Callahan et al., 2016), and followed the work-flow suggested by the developers. The detailed protocol is described in the SI.

2.5. Chemical analysis

Metals, ions (i.e., dissolved cations and anions), nutrients, and dissolved organic carbon were measured as described in Ju et al. (2019) using high-resolution inductively coupled plasma mass spectrometry, ion chromatography, flow-injection analysis, and total organic carbon analyzer, respectively, as described in the SI. Dissolved micropollutants (i.e., pharmaceuticals, antibiotics) were measured as described in Ju et al. (2019) using liquid chromatography triple quad mass spectrometry with electrospray ionization in the SI. Total dried solid (TS) were measured in sediment samples according to standard methods (APHA-AWWA-WPCF., 1981).

2.6. Estimating the dilution effect on downstream levels of resistance determinants

Under continuous discharge and after complete horizontal and vertical mixing, the discharged load of a conservative tracer (e.g., sodium) entering the river through EF is expected to be conserved along the river continuum. Under this assumption, any change in the concentration of the conservative tracer would be due to dilution effects by additional water inflows (i.e., groundwater and/or tributary inputs) and additional inputs of the tracer with these inflows. We used sodium and two micropollutants as conservative tracers (i.e., 4/5-methylbenzotriazole, carbamazepine) because these substances had high concentrations in EF compared to the US river and are known to not substantially degrade or adsorb in the river system. The rationale for selecting the conservative tracers is described in more detail in the SI.

Starting with these mass conservation assumptions, under steady state conditions, the dilution parameter (*DP*, the ratio between external water inflow and streamflow at the downstream section

between any two points A and B along a river stretch) can be estimated from a ratio of tracer concentrations according to Eq.1:

$$DP_{A\rightarrow B} = (C_B - C_A)/(\bar{C} - C_A) \qquad \textbf{(Eq.1),}$$

where A indicates an upstream location; B denotes a downstream location; C indicates the concentration of a tracer; $\bar{C}$ denotes the average concentration of the tracer in the external inflow between A and B.

The derivation of Eq.1 is schematized in Fig. 1, and also described in detail in the SI. Concentrations C_A or C_B were measured directly for all compounds, $\bar{C}$ was estimated according to the following equation (Eq.2) for sodium (the values shown in Dataset S9.3) and assumed to be 0 for 4/5-methylbenzotriazole and carbamazepine. In short, the difference in sodium loadings (mass per time) between the point of EF discharged and the downstream point where gauging stations were located (D8 for VIL; D4 for MUE) was divided by the quantity of additional water inflows:

$$\overline{Na_{in}} \approx \frac{Na_{D8\ or\ D4} \times Q_{D8\ or\ D4} - (Na_{US} \times Q_{US} + Na_{EF} \times Q_{EF})}{Q_{D8\ or\ D4} - (Q_{US} + Q_{EF})} \qquad (\textbf{Eq. 2})$$

Where, $Na_{D8\ or\ D4}$ denotes the sodium concentration measured at D8 or D4; $Q_{US,EF,D8\ or\ D4}$ indicates the river flow quantity or wastewater effluent discharge (volume per time) at US, EF, D8 or D4.

The value *DP* should be the same for all conservative tracers. To test our hypothesis that "the short distance dynamics of resistance determinants is largely governed by dilution effects", we calculated *DP* over short distance (i.e., $DP_{D1\rightarrow D2}$ for VIL, and $DP_{D1\rightarrow D3}$ for MUE), and compared *DP* values over the same distance for resistance determinants (*sul1*, *intI1*, *ermB*, *tetW*, and CLR/TET resistant bacteria) with values for the conservative tracers. Higher *DP* values for resistance determinants would indicate a lower than expected concentration in the downstream and thus removal.

The expected downstream concentrations of resistance determinants considering dilution as a main driver can be calculated using the $DP_{A\rightarrow B}$ of conservative tracers according to the following relationship:

$$C_{resist-B} = C_{resist-A} - C_{resist-A} \times DP_{A\rightarrow B\ (\text{for X})} \qquad \textbf{(Eq.3),}$$

where $C_{resist-A}$ indicates the concentration of a resistance determinant at an upstream location A; $C_{resist-B}$ denotes the concentration of a resistance determinant at a downstream location B; $DP_{A\rightarrow B(\text{for X})}$ indicates the *DP* of a conservative tracer (X) between A and B

Eq.3 assumes that resistance determinants behave conservatively over the studied distances and that there are no significant inputs of resistance determinants from the diluting water inflows (i.e., $\bar{C}$ for resistance determinants ≈ 0 in Eq.1). Therefore, deviations from measured to predicted values can indicate violation of these assumptions. We calculated the predicted concentration by dilution effects for each resistance determinant under these assumptions for all downstream sections of the rivers.

2.7. Estimating the river discharge over downstream distance

The river discharge (*Q*) at was estimated for several downstream locations where there were not gauging stations. The estimated *Q* values were used when calculating loadings of chemical and resistance indicators over downstream distance. The Q_{EF} values were obtained from each WWTP, and Q_{US} values were either obtained from gauging station (for VIL), or calculated as shown in Eq.8 in the SI (for MUE).

$$\boldsymbol{Q_{D1} = Q_{US} + Q_{EF},}$$

$$\textbf{If n} > 2, \boldsymbol{Q_{D(n)} = Q_{D(n-1)} + Q_{D(n-1)} \times \frac{DP_{D(n-1)\rightarrow n}}{1-DP_{D(n-1)\rightarrow n}}} \qquad \textbf{(Eq.4)}$$

Where, $Q_{D(n)}$ indicates the river discharge (*Q*) at the downstream location D(n) (2 ≤ n ≤ 8); $DP_{D(n-1)\rightarrow n}$ denotes the dilution parameter between D(n-1) and D(n).

3. Results and Discussion

3.1. Upstream water quality and WWTP effluent

In agreement with the criteria for site selection, the levels of *intI1* and target ARGs upstream of the WWTP were generally low, except for the MUE2 campaign where we observed elevated upstream levels of *ermB*, and *intl1* (Fig. 2; Fig. S4 and S5 in the SI). Chemical water quality likewise did not suggest significant pollution inputs from either tributaries or upstream locations for either VIL or

MUE as most micropollutants were below the limit of quantification (Dataset S10~11). Certain micropollutants (e.g., 4/5-methylbenzotriazole, benzotriazole, and diclofenac) were sporadically detected in very low quantities. For cultivable multi-resistant bacteria (Fig. 3), especially CLR/TET resistance, relatively high upstream values were observed in VIL2 and MUE2-US samples. These findings indicate that while there is no indication of significant upstream pollution, some pollution, probably from periodical urban or agricultural activities may affect the river. There was a settlement upstream of the WWTP and livestock farming activities (i.e., pastures and meadows for livestock) in the catchments, including the upstream sections, in both sites (BAFU, 2013). While we assume surface runoff from the agricultural sites to be minimal as our samplings were performed under dry-weather condition, it cannot be ruled out that some inputs from agricultural activity occasionally affected the river. Further investigations into the nature of these transient microbial contaminations were not carried out in this project, but future work could employ microbial source tracking or microbial fingerprinting approaches to determine their sources.

The effluent from both WWTPs contained considerable levels of pollutants. For instance, effluent concentrations were higher than the upstream levels by approximately 1 order of magnitude for sodium, 1~2 order of magnitude for ARGs and *int*I1, and more than 2 order of magnitude for micropollutants (Datasets S8~11). These results are in line with previous results from a large-scale research on micropollutants in Swiss streams (Stamm et al., 2016).

3.2. Short range fate of antibiotic resistance determinants in the downstream river

Focusing on the immediate impact of the WWTP effluents (US versus D1 to D3 sites), there were significant impacts of WWTP effluents on the receiving rivers in both VIL and MUE. The estimated proportions of EF in the downstream receiving waters (D1) estimated by conductivity were 10.5 ~ 35.9 % for VIL1~4, and 33.0 ~ 38.0 % for MUE1~3 (Dataset S9.2). Accordingly, significant increases of *sul1*, *intI1*, *tetW* and *int*I1 as quantified by qPCR were observed at D1 compared to US ($p<0.01$ paired t-test; Fig. 2 & Fig. S5). However, the measured levels of these antibiotic resistance indicator genes rapidly decreased nearly to close to upstream levels over 2.5 and 2 km downstream distance (D2 or D3 locations) in VIL and MUE, respectively.

286 The same dynamic was also observed for multi-resistant bacteria (Fig. 3), especially CLR/TET
287 resistance. SMX/TMP/TET resistance was often below the limit of detection (5.0 CFU/mL), but
288 clearly exceeded it in the D1 samples and was thus also higher there than further downstream (from
289 D2 on).

290 Several processes may contribute to the observed decrease of resistance determinants, including
291 dilution by additional water inflows via groundwater and/or tributary inputs, biological deterioration
292 (e.g. cell death or dormancy due to exposure to sunlight, lower ambient temperature, predation, etc.),
293 and cell sedimentation.

294 *3.3. Dilution effects strongly affect short distance dynamics of effluent resistance determinants*

295 To determine the importance of dilution effects, we compared DP calculated over a short distance
296 (D1 to D2 for VIL; D1 to D3 for MUE) downstream of the WWTP discharge point ($DP_{D1\rightarrow 2}$ for VIL;
297 $DP_{D1\rightarrow 3}$ for MUE) from conservative chemical tracer concentrations (e.g., sodium, 4/5-
298 methylbenzotriazole, carbamazepine) as well as ARG and *intI1* levels (Fig. 4). The average
299 $DP_{D1\rightarrow 2\ or\ 3}$ of the target antibiotic resistance indicator genes levels were always higher than for
300 conservative tracers, indicating possible removal mechanisms at play. However, according to the
301 paired t-test under the null-hypothesis of "No significant differences of dilution parameters between
302 different pairs of bio- and conservative indicators", only the differences between *sul1* and *tetW* versus
303 sodium were significant at $p<0.05$ (p-adjusted using Benjamini-Hochberg method) (Fig. 4),
304 confirming non-conservative behavior and additional removal mechanisms. As $DP_{D1\rightarrow 2\ or\ 3}$ for
305 sodium took up a large portion of the values for *sul1* and *tetW* (i. e. $DP(Na^+)_{D1\rightarrow 2\ or\ 3} =$
306 $0.72 \sim 0.92 \times DP(sul1)_{D1\rightarrow 2\ or\ 3}$ and $0.59 \sim 0.96 \times DP(tetW)_{D1\rightarrow 2\ or\ 3}$), the dilution effects
307 quantified by sodium nonetheless explain the majority of the concentration decrease for these
308 parameters. This result implies that the observed rapid decrease in the downstream levels of
309 wastewater-origin resistance determinants immediately downstream of the WWTPs was mainly
310 governed by dilution in the studied systems. Dilution effects thus need to be carefully considered in
311 studies of the environmental fate of resistance determinants, and loadings instead of concentrations
312 need to be determined to accurately assess environmental fate.

3.4. Additional source/sink effects become apparent over longer downstream distances

We hypothesized that additional source/sink mechanisms affect the downstream behaviors of antibiotic resistance indicator genes over longer distances. To analyze this in more detail the daily loading (copies/day) for the target ARGs and *intI1* at the point of discharge (as the sum of upstream and EF loadings), and for short (D2 for VIL; D3 for MUE) and far downstream distances (D8) were calculated by multiplying resistance levels (copies/m^3) with the discharge (m^3/day) at each location and then compared. The discharge was either obtained directly from nearby gauging stations, or estimated under consideration of the EF discharge (m^3/day) using sodium as an indicator, and according to the equation derived under the mass-conservation assumption (Eq.4).

The downstream behaviors of the target antibiotic resistance indicator genes varied by indicator and also by sampling campaign. For instance, the load decrease from wastewater discharge (US+ EF) to the furthest downstream point (~ D8) was pronounced and consistent for *ermB* and *tetW* in all the samplings (Fig. 4b & c). The average load reduction was 81±17 % for *ermB*, and 70±15 % for *tetW* over 13.7 km distance in VIL1~4; 95±5 % for *ermB*, and 96±2 % for *tetW* over the 6.8 km distance in MUE1~3 (Dataset S13). In contrast, the downstream behavior of the *sul1* loadings was inconsistent between sampling campaigns. A pronounced decrease over distance (64-94 % at D8) was seen in VIL1 ~ 4, but little reduction over distance in MUE1 ~ 2 (7 and 29% at D8), and a strong increase in MUE3. The downstream fate of *intI1* was also variable, for instance, as *intI1* loads did not decrease and in some instances even increased .

To further analyze if there is a break point where *sul1* and *intI1* start to deviate from conservative behavior, we calculated the predicted levels of resistance determinants over the whole study distance considering dilution as a major driver using Eq.3 (Dataset S8). The predictions based on three different conservative tracers are visualized for *sul1* and *intI1* in Fig. 2. In VIL, measured levels are always below predictions, except for *intl1* in VIL2. For MUE, in contrast, we see measured values exceeding predicted values in several instances, for *intl1* even for most downstream locations. In MUE3 where the pronounced increase of *sul1* loading was observed between D5 and D8, the level of *sul1* started to exceed predicted values between 5 ~ 6.8 km distance. The concentration of *intI1*

increased also very rapidly between 5 ~ 6.8 km downstream distance in MUE3, which indicates either a pronounced proliferation or a non-point source of *sul1* and *intI1* in this stretch of the river. We will discuss potential mechanisms (e.g., biological drivers, on-site selection, additional anthropogenic source input, and surface sediment inputs) in section 3.10.

A number of mechanisms may contribute to the generally observed removal: Sedimentation (especially of cell aggregates or flocs) and cell death by predation, UV light, or vaious other environmental conditions unfavorable to wastewater bacteria. With the available data we are not able to determine the contribution of various mechanisms. Future studies could investigate the persistence of resistant bacteria or molecular resistance markers in micro- or mesocosm experiments or in a turbulent flow system mimicking natural streams to answer such questions. Modeling transport and sedimentation of wastewater-origin particles using the information on particle size, mass, and flow characteristics could provide information on the importance of sedimentation.

3.5. WWTP effluent affects the downstream riverine resistome

To obtain a broader view of the river antibiotic resistome we retrieved the ARG content of metagenomes obtained for selected sampling campaigns (VIL1, MUE2, and MUE3). Overall, 65 ARG subtypes were identified, 49 of them occurred in both upstream and downstream river samples (Dataset S7). The antibiotic resistome in the receiving water closest to the discharge point (D1) was clearly influenced by the input from EF. For instance, a total 28 out of 36 ARG subtypes found in D1 were also detected in EF (B, C, F, G in Fig. 5a) while 16 of these were not observed US (B, C in Fig. 5a). The 16 EF-derived resistance genes confer resistance to the following antibiotic classes: 1 x aminoglycoside, 4 x beta-lactam antibiotics, 1 x chloramphenicol, 2 x macrolide, 1 x quinolone, 4 x tetracycline; 3 subtypes were multidrug resistance genes. Of these 16 genes, 14 were no longer detected in the far downstream (6.8 ~ 13.7 km downstream distance in MUE and VIL, designated as D_Far in Fig. 5a). This is in agreement with the results for qPCR enumeration of ARGs and *intI1* reported above and implies that the majority of ARGs that occurred exclusively in EF do not persist at detectable levels for a long distance in rivers where significant amounts of additional water inflows and additional removal mechanisms are expected.

3.6. Diversity of the river resistome and microbiome along the river continuum

We analyzed the alpha-, and beta-diversity of river resistomes and microbiomes as another way to observe potential effects of the WWTP discharge and to see if the dynamics in the resistome are strongly correlated with the microbial community, as noted e.g. for changes observed during wastewater treatment (Ju et al., 2019). As expected, Shannon alpha-diversity of ARGs was higher in EF compared to US samples by 20.2 ~ 225.4 % (Fig. 5b). Accordingly, the impact of the EF resistome was observed, especially for VIL1 and MUE3, as an increase in ARG diversity in river water at the D1 sites. The ARG alpha-diversity decreased downstream in all sampling campaigns. However, for the microbial community as represented by amplicon sequence variants (ASVs), alpha-diversity was not consistently higher in EF versus US, and consequently also did not change significantly from US to D1 and did not consistently decrease downstream (Fig. 5b). This indicates that the downstream dynamics of the overall microbial community and the antibiotic resistome were decoupled.

Similar conclusions were obtained from beta-diversity analysis. Procrustes analysis between microbial communities and antibiotic resistome (Fig. 5c) revealed a strong structural correlation between microbial communities and antibiotic resistome only when the most strongly effluent-affected sites (D1) were considered ($p = 0.002$; Fig. 5c). When the D1 samples were excluded, the correlation was barely significant ($p = 0.04$; Fig. 5d), indicating that the structural correlation between microbial communities and resistome largely resulted from the impacts of WWTPs on D1. The weak structural correlation in less impacted waters suggests a lack of strong drivers, such as selective pressures or the influx of external ARB.

3.7. Resistome analysis confirms effluent effect and abatement

To quantitatively investigate the dynamics of the resistome along the river continuum in more detail, the seven most prevalent ARG subtypes that appeared in more than 9 out of 18 samples were chosen for detailed analysis. We calculated the proportion of each gene in this set based on relative abundance (GPM, gene per million) (Fig. 6a). The proportions (%) of each of six genes (*aph(3")-I*, *aph(6)-I*, *mexT*, *tetQ*, *aadA*, and *sul1*) to the whole seven genes were lower in US than in EF and D1 (Fig. 6a). The *bacA* gene, in contrast, comprised a large proportion in US (i.e., up to 83 % in

VIL1:US), D5 and D8 (i.e., up to 93 % in VIL1:D8) than in EF and D1. It was therefore excluded from the plots of relative and cumulative abundance of the assembled ARG in the metagenome in Fig. 6b, and their individual and cumulative environmental level (Gene per liter) in Fig. 6c. Both relative and absolute abundances showed a similar pattern – the abundances of the selected ARG were higher in EF and D1 compared to US (Fig. 6b). The abundances of those six genes decreased along the downstream locations except for *sul1* in the far downstream location (D8) in the MUE3 campaign (Fig. 6b). This analysis confirmed a quantitative effect of the effluent on the abundance of prevalent resistance genes in the river resistome, and suggests additional candidates for tracking anthropogenic sources of resistance in future studies (*aph(3'')-I*, *aph(6)-I*, *mexT*, *tetQ*, *aadA*, and *sul1*) that may be useful for tracking resistance determinants from wastewater. Several of these genes have been used as resistance indicators for environmental samples mainly in combination with culture-dependent approaches (Rizzo et al., 2013;Zhang et al., 2009), but much less frequently with culture-independent approaches (Rizzo et al., 2013;Sharma et al., 2016).

3.8. Contig analysis indicates ARG co-location

We hypothesized that the similar dynamics of some ARGs could derive from their co-location in the same host or presence on the same genetic elements, so we analyzed their loci within the contigs. *aph(3'')-I* and *aph(6)-I* were indeed found to be located on the same contig in many samples from VIL1 (EF, D1, D2, and D5) (Fig. 7a), which may explain the strongly similar dynamics between *aph(3'')-I* and *aph(6)-I* GPM in VIL1 with $R^2 = 0.98$ ($p < 0.001$) (Fig. 6b). Another case of co-location between ARGs was observed between *sul1* and *aadA*. The contigs containing those two genes were found in D1 in VIL1, and EF in MUE3 (Fig. S13c; Fig. 7b). Unlike *aph(3'')-I* and *aph(6)-I* in VIL1 however, the dynamics of *sul1* was not similar to that of *aadA* especially in D5 and D8 in MUE3 where high abundance of *sul1* was observed while no *aadA* was assembled (Fig. 6b). In agreement with this observation, the *sul1*-containing contigs retrieved from D5 and D8 in MUE3 did not contain *aadA*, and this was the only type that was identified in those samples (Fig. 7b). This indicates a significant shift of the bacterial populations that yielded *sul1*-containing contigs between D1 and D3 in MUE3, with wastewater-derived contigs containing *sul1 – aadA* pairs not persisting.

3.9. bacA, an ARG with high natural prevalence in environmental bacteria

Unlike the other prevalent genes, the proportion of *bacA* (also known as *UppP*, undecaprenyl-diphosphate or -pyrophosphate phosphatase) was greatest in US samples, and was also abundant in many downstream locations (except D5 and D8 in MUE3 where *sul1* occupied the largest proportion) (Fig. 6a). In order to further assess whether *bacA* was intrinsic in our river water samples, we identified potential hosts by assigning taxonomy to the metagenome-assembled contigs using a combination of methods. The contigs for which all three methods agreed at the genus level are shown in Table 1. The four genera identified as potential hosts of *bacA* contigs derived from less-disturbed freshwater samples (US, D2, D5, D8) were *Pseudomonas*, *Acidovorax*, *Limnohabitans*, and *Aeromonas*. Among them, *Pseudomonas*, *Acidovorax*, and *Limnohabitans* are typical inhabitants of freshwater and soil environments (Peix et al., 2009;Willems, 2014). However, considering that the proportions of *bacA* in the contigs to the total *bacA* in the sample in terms of reads per kilobase (RPK) were low for river water samples (except for D8 in MUE3), we assume that homologues of *bacA* could be present in many other environmental bacteria. Thus, our data suggests that *bacA* is probably unsuited for tracking anthropogenic sources of antibiotic resistance.

3.10. Exploring the potential reasons for rapid increase of sul1 in far downstream locations in MUE3

Both qPCR-based, and metegenomic analysis confirmed that *sul1* and *intI1* increased in the downstream of MUE, and especially strongly in one of the sampling campaigns (i.e., MUE3) between 5.0 – 6.8 km downstream distance (Fig. 2b & 6).

To figure out if there was a biological driver for this unexpected increase of *sul1* and *intI1*, we first characterized the *sul1*-containing contigs. The *sul1* gene is known to be highly mobilized and is often associated with *intI1* (Gillings et al., 2008;Gillings et al., 2015). Indeed, all contigs containing *sul1* associated with *intI1* retrieved from the river (D3 ~ 8) were homologs of a single dominant type that appeared to be also plasmid-associated as it contained the plasmid-associated gene *parA* (Davis et al., 1992). We could unsurprisingly not obtain a meaningful taxonomic assignment for these sequences. It could thus not be demonstrated whether the downstream increase in MUE3 was due to an increase in

448 an EF-derived or an environmental organism or from a local contamination source. Further
449 information could be obtained in future studies by isolation of construction of metagenome-assembled
450 genomes. We therefore turned to chemical indicators to further study the potential for local or non-
451 point source contamination as an explanation.

452 To evaluate non-point-source inputs of pollutants, we chose to evaluate a few micropollutants that
453 may serve as indicators of contamination. Sulfamethazine (also known as sulfadimidine) is used in pig
454 husbandry (Stoob et al., 2007), and mecoprop is a weed control agent used primarily in urban settings
455 in Switzerland (Wittmer et al., 2010). It was assumed that the levels of these pollutants in downstream
456 locations would increase or be persistently high if a pronounced agricultural or urban surface runoff
457 existed, which could accompany resistance genes and bacteria potentially existing in agricultural or
458 urban area. In VIL, sulfamethazine was below detection (LOD ~ 0.8 ng/L) in all samples except one
459 US sample, while in MUE there appeared to be a source in WWTP effluent especially during the
460 MUE3 campaign, but the compound was not observed to increase in downstream locations. The
461 concentrations and downstream dynamics of mecoprop varied between campaigns (Fig. S8). For
462 VIL1~3 and MUE1 and MUE3 mecoprop concentrations were low (< 60 ng/L), while there seemed to
463 be a strong, effluent-associated input for MUE2 and concentrations remained high further downstream
464 (> 200 ng/L). A slight increase in the downstream range > 5km observed in MUE2 and between 1.0 –
465 2.0 km in MUE3 may be due to fluctuating input of mecoprop from the WWTP effluent.
466 Concentrations did not further increase in the far downstream locations (D8) where the sudden
467 increase of *sul1* and *intI1* was observed (Fig. S8). Based on these, but also at the other analyzed
468 micropopllutants, we found no evidence for significant downstream contamination sources. However,
469 these chemical indicators are not conclusive, as the analyzed compounds were not a comprehensive
470 selection to trace non-point sources in the downstream river section (e.g., from manure or pesticide
471 applications, although these are not very likely at dry-weather conditions). So while we found no
472 evidence for such contamination we can also not conclusively rule them out as an explanation for the
473 marked *sul1* and *intl1* increase observed for MUE3.

Finally, the potential for in-situ resistance selection in the water was assessed using the concentration of antibiotics and metals in downstream locations in MUE3. Sulfamethoxazole and its derivative (N4-acetylsulfamethoxazole) were the antibiotics with the highest concentration among all the antibiotics analyzed, but downstream concentration (sulfamethoxazole in the range of 33 to 95 ng/L in MUE3) remained far below the published PNEC for resistance selection (e.g., 16,000 ng/L) (Bengtsson-Palme and Larsson, 2016). The concentration of trimethoprim, which is usually prescribed together with sulfamethoxazole, was also much lower than its PNEC for resistance selection (e.g. 500 ng/L) (Bengtsson-Palme and Larsson, 2016). Even though the vast majority of metals analyzed in this study showed below limit of quantification or below their estimated minimum co-selective concentrations for dissolved metals in water ($MCC_{waterDC}$), the concentrations of two metals (i.e., copper and nickel) were higher than their $MCC_{waterDC}$ (1.5 µg/L for Cu, and 0.29 µg/L for Ni) (Seiler and Berendonk, 2012). However, $MCC_{waterDC}$ is a predictive value and actual selective levels could be higher, also their levels in far downstream locations (D8) in MUE3 were not specifically higher than at other locations within the same sampling campaign, nor at the same locations than in other samplings where the increase of *sul1* or *intl1* was less pronounced (Dataset S12). Furthermore, we could not observe the co-localization between *sul1* and any other genes potentially conferring Cu, Ni or any other metal resistance based on contig-based co-localization search in MUE (Fig. 7b). Overall no convincing evidence for in-situ resistance co-selection by Cu and Ni as an explanation for the downstream increase of *sul1* and *intl1* was found. We further note that the estimated river retention time per km was relatively short (i.e., 51.4 and 49.5 mins/km for VIL and MUE, respectively, Dataset S4), which makes the likelihood of in-situ resistance selection in the water even less plausible.

As a final possible explanation we considered the possibility of cell migration from other river compartments to the water. According to qPCR enumeration of ARGs and *intI1* in surface sediments, we did not observe the increase of *sul1* and *intI1* levels in D5 and D8 in MUE3 in terms of either absolute and relative abundance (Fig. S6). Furthermore, the relative abundance of *sul1* and *intI1* in sediment was generally similar to, or lower than the values for water in MUE1~3 (Fig. S7). If sediment resuspension was a major source for aquatic *sul1* and *intI1* elevation in MUE3 D8, relative abundances of *sul1* and *intI1* in water samples would be expected to remain unchanged or to drop.

While we could not completely exclude the possibility of contribution of sediment resuspension, we assume that there could be other sources (e.g., stream biofilms) where *sul1* and *intI1* were selectively enriched in terms of both absolute and relative abundance. Considering the downstream levels of both resistance determinants and nutrients remained relatively high in MUE due to high EF inputs and low downstream dilution effects, especially in the third campaign (Fig. S5 & Dataset S9), in-system growth is a plausible hypothesis.

The reason for and the nature of the striking increase of *sul1* and *intI1* (but not of *tetW*, *ermB*, bla_{CTX}), during the MUE3 campaign thus remains open and would require further study to resolve. What the observation shows unambiguously, is that unexpected and perhaps not directly anthropogenic contamination-driven increases of ARGs are possible. As in particular *sul1* and *intI*1 are commonly applied for tracking anthropogenic sources of ARG in the environment (Berendonk et al., 2015;Gillings et al., 2015), we caution that monitoring strategies should employ a multi-target strategy to be robust.

4. Conclusions

- Downstream levels of antibiotic resistance determinants decreased rapidly over 2.0 – 2.5 km distance due to dilution effects and decay over longer distance due to other removal mechanisms. This would suggest that public exposure to wastewater-origin antibiotic resistance might be most acute only over short distances (few kilometers) from points of discharge, especially if a pronounced input of additional water inflows exists.
- We also observed at least one instance where *sul1* and *intI1* dynamically increased in the river, without being able to establish any link to a local anthropogenic contamination. Other river compartments where in-system growth of biomass could take place (i.e., stream biofilms) could be included as a monitoring target in future studies.
- Metagenomics-based resistome analysis yielded consistent conclusions with qPCR analysis of select targets (e.g., *sul1*) and also identified promising targets for future monitoring of anthropogenic sources of antibiotic resistance (e.g., *aph(3'')-I*, *aph(6)-I*, *mexT*, *tetQ*, and *aadA*). In general metagenomics, qPCR and cultivation-based assays yielded consistent trends. Public health

advice could be based on quantifying indicator genes or technically simpler cultivation-based indicators.

- A weak structural correlation between resistome and microbiome, and low levels of (co)selective agents revealed a lack of driving forces in less-disturbed river waters (downstream over 3 km distance, plus upstream locations).
- We showed that contig-based taxonomic assignment and analysis of the genetic neighborhood of assembled ARG can reveal important, if limited, additional information about shifts in ARG host identities, mobilization, and co-localisation of ARG that would otherwise remain hidden.

Acknowledgements

This study was financially supported by the Swiss National Science Foundation under the National research Programme NRP72 "Antimicrobial resistance", grant 167116 and the U.S. National Science Foundation Partnership in International Research and Education Award 1545756. Dr. Feng Ju received financial support from the NSFC Young Scientists Fund via Project 51908467. We would like to thank to Prof. Alex Hall from ETH Zurich, and Prof. Célia Manaia from Catholic University of Portugal for providing expert advice. We acknowledge technical support from current or former Eawag colleagues, including Patrick Kathriner, Nina Studhalter, Melea Brunner, Lian Yang, Irene Brunner, and Thomas Rüttimann. We also appreciate the support of the Genetic Diversity Centre (GDC) at ETH Zurich for providing us with bioinformatics consultation as well as access to computational resources. We owe special thanks to Dr. Cintia Ramón Casañas from Eawag for giving us her valuable insights on hydrology, and to Heinz Singer of Eawag for help with chemical analysis. Finally, we want to thank to all the personnel from the WWTPs at Villeret and Münchwilen for access to their facilities, help during on-site samplings, and sharing information.

Data availability

The raw sequencing data both for metagenome and 16S rRNA amplicon sequencing are available at the European Nucleotide Archive under the project ID – ERP123247. All the other research datasets (including additional minor datasets that were not shown in this manuscript, and R codes) will be

available at the Eawag Research Data Institutional Collection (https://opendata.eawag.ch/) upon acceptance of this manuscript to the journal.

Declaration of competing interest

The authors declare that there are not any competing interests regarding any issues related with this study.

Author contributions

J.L and H.B designed this study, and participated in all stages of the work as main authors. All authors provided feedback and inputs throughout field/laboratory works or manuscript writing stages. F.J, A.M.M, and K.B participated in field/laboratory works and biological data analysis. C.S.M, A.M, participated in designing the study, and performed chemical analysis of micropollutants. M.D.M, and F.F and helped design hydrological measurements and experiments and together with C.S. helped with analysis and interpretation of hydrological data. P.V, A.P, and C.S provided important input and a critical review of the manuscript.

References

Albertsen, M., Hugenholtz, P., Skarshewski, A., Nielsen, K.L., Tyson, G.W. and Nielsen, P.H. (2013) Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. Nature biotechnology 31(6), 533-538, 10.1038/nbt.2579.

Andersson, D.I. and Hughes, D. (2012) Evolution of antibiotic resistance at non-lethal drug concentrations. Drug resistance updates 15(3), 162-172, 10.1016/j.drup.2012.03.005.

APHA-AWWA-WPCF. (1981) Standard methods for the examination of water and wastewater, APHA American Public Health Association.

BAFU (2013) Land use statistics 2004/09 based on the standard nomenclature NOAS04, Bundesamt für Umwelt (BAFU), CH-3003 Bern.

Bengtsson-Palme, J. and Larsson, D.G. (2016) Concentrations of antibiotics predicted to select for resistant bacteria: Proposed limits for environmental regulation. Environment international 86, 140-149, 10.1016/j.envint.2015.10.015.

582 Berendonk, T.U., Manaia, C.M., Merlin, C., Fatta-Kassinos, D., Cytryn, E., Walsh, F., Burgmann, H.,
583 Sorum, H., Norstrom, M., Pons, M.N., Kreuzinger, N., Huovinen, P., Stefani, S., Schwartz, T.,
584 Kisand, V., Baquero, F. and Martinez, J.L. (2015) Tackling antibiotic resistance: the environmental
585 framework. Nature reviews microbiology 13(5), 310-317, 10.1038/nrmicro3439.
586 Bürgmann, H., Frigon, D., H Gaze, W., M Manaia, C., Pruden, A., Singer, A.C., F Smets, B. and
587 Zhang, T. (2018) Water and sanitation: an essential battlefront in the war on antimicrobial
588 resistance. FEMS microbiology ecology 94(9), fiy101, 10.1093/femsec/fiy101.
589 Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J. and Holmes, S.P. (2016)
590 DADA2: High-resolution sample inference from Illumina amplicon data. Nature methods 13(7),
591 581-583, 10.1038/nmeth.3869.
592 Cassini, A., Hogberg, L.D., Plachouras, D., Quattrocchi, A., Hoxha, A., Simonsen, G.S., Colomb-
593 Cotinat, M., Kretzschmar, M.E., Devleesschauwer, B., Cecchini, M., Ouakrim, D.A., Oliveira,
594 T.C., Struelens, M.J., Suetens, C., Monnet, D.L. and Burden of, A.M.R.C.G. (2019) Attributable
595 deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in
596 the EU and the European Economic Area in 2015: a population-level modelling analysis. The
597 Lancet infectious diseases 19(1), 56-66, 10.1016/S1473-3099(18)30605-4.
598 Cockerill, F., Patel, J., Alder, J., Bradford, P., Dudley, M. and Eliopoulos, G. (2013) Performance
599 standards for antimicrobial susceptibility testing: twenty-third informational supplement; M100-
600 S23. Wayne, PA: CLSI.
601 Czekalski, N., Berthold, T., Caucci, S., Egli, A. and Burgmann, H. (2012) Increased levels of
602 multiresistant bacteria and resistance genes after wastewater treatment and their dissemination into
603 Lake Geneva, Switzerland. Frontiers in microbiology 3, 106, 10.3389/fmicb.2012.00106.
604 Czekalski, N., Diez, E.G. and Burgmann, H. (2014) Wastewater as a point source of antibiotic-
605 resistance genes in the sediment of a freshwater lake. The ISME journal 8(7), 1381-1390,
606 10.1038/ismej.2014.8.
607 Davis, M.A., Martin, K.A. and Austin, S.J. (1992) Biochemical activities of the parA partition protein
608 of the P1 plasmid. Molecular microbiology 6(9), 1141-1147, 10.1111/j.1365-2958.1992.tb01552.x.

609 Eawag (2014): WWTP of Switzerland: Revised on the basis of the project: Maurer M. und Herlyn A.
610 (2007) Status, costs and in-vestment needs of Swiss wastewater disposal. Eawag/Bafu report (in
611 German).
612 Gillings, M., Boucher, Y., Labbate, M., Holmes, A., Krishnan, S., Holley, M. and Stokes, H.W. (2008)
613 The evolution of class 1 integrons and the rise of antibiotic resistance. Journal of bacteriology
614 190(14), 5095-5100, 10.1128/JB.00152-08.
615 Gillings, M.R., Gaze, W.H., Pruden, A., Smalla, K., Tiedje, J.M. and Zhu, Y.G. (2015) Using the class
616 1 integron-integrase gene as a proxy for anthropogenic pollution. The ISME journal 9(6), 1269-
617 1279, 10.1038/ismej.2014.226.
618 Hyatt, D., Chen, G.L., Locascio, P.F., Land, M.L., Larimer, F.W. and Hauser, L.J. (2010) Prodigal:
619 prokaryotic gene recognition and translation initiation site identification. BMC bioinformatics
620 11(1), 119, 10.1186/1471-2105-11-119.
621 Ju, F., Beck, K., Yin, X., Maccagnan, A., McArdell, C.S., Singer, H.P., Johnson, D.R., Zhang, T. and
622 Burgmann, H. (2019) Wastewater treatment plant resistomes are shaped by bacterial composition,
623 genetic exchange, and upregulated expression in the effluent microbiomes. The ISME journal
624 13(2), 346-360, 10.1038/s41396-018-0277-8.
625 Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. Nature methods
626 9(4), 357-359, 10.1038/nmeth.1923.
627 Lee, K., Kim, D.W., Lee, D.H., Kim, Y.S., Bu, J.H., Cha, J.H., Thawng, C.N., Hwang, E.M., Seong,
628 H.J., Sul, W.J., Wellington, E.M.H., Quince, C. and Cha, C.J. (2020) Mobile resistome of human
629 gut and pathogen drives anthropogenic bloom of antibiotic resistance. Microbiome 8(1), 2,
630 10.1186/s40168-019-0774-7.
631 Li, D., Liu, C.M., Luo, R., Sadakane, K. and Lam, T.W. (2015) MEGAHIT: an ultra-fast single-node
632 solution for large and complex metagenomics assembly via succinct de Bruijn graph.
633 Bioinformatics 31(10), 1674-1676, 10.1093/bioinformatics/btv033.
634 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin,
635 R. and Genome Project Data Processing, S. (2009) The Sequence Alignment/Map format and
636 SAMtools. Bioinformatics 25(16), 2078-2079, 10.1093/bioinformatics/btp352.

637 Li, J., Cheng, W., Xu, L., Jiao, Y., Baig, S.A. and Chen, H. (2016) Occurrence and removal of
638 antibiotics and the corresponding resistance genes in wastewater treatment plants: effluents'
639 influence to downstream water environment. Environmental science and pollution research 23(7),
640 6826-6835, 10.1007/s11356-015-5916-2.
641 Mansfeldt, C., Deiner, K., Machler, E., Fenner, K., Eggen, R.I.L., Stamm, C., Schonenberger, U.,
642 Walser, J.C. and Altermatt, F. (2020) Microbial community shifts in streams receiving treated
643 wastewater effluent. Science of the total environment 709, 135727,
644 10.1016/j.scitotenv.2019.135727.
645 Menzel, P., Ng, K.L. and Krogh, A. (2016) Fast and sensitive taxonomic classification for
646 metagenomics with Kaiju. Nature communications 7(1), 11257, 10.1038/ncomms11257.
647 O'Neill, J. (2014) 'Review on Antimicrobial Resistance. Antimicrobial Resistance: Tackling a Crisis
648 for the Health and Wealth of Nations. 2014. The Wellcome Trust and UK Government.
649 Peix, A., Ramirez-Bahena, M.H. and Velazquez, E. (2009) Historical evolution and current status of
650 the taxonomy of genus Pseudomonas. Infection, genetics and evolution 9(6), 1132-1147,
651 10.1016/j.meegid.2009.08.001.
652 Rizzo, L., Manaia, C., Merlin, C., Schwartz, T., Dagot, C., Ploy, M.C., Michael, I. and Fatta-Kassinos,
653 D. (2013) Urban wastewater treatment plants as hotspots for antibiotic resistant bacteria and genes
654 spread into the environment: A review. Science of the total environment 447, 345-360,
655 10.1016/j.scitotenv.2013.01.032.
656 Sabri, N., Schmitt, H., Van der Zaan, B., Gerritsen, H., Zuidema, T., Rijnaarts, H. and Langenhoff, A.
657 (2018) Prevalence of antibiotics and antibiotic resistance genes in a wastewater effluent-receiving
658 river in the Netherlands. Journal of environmental chemical engineering, 102245,
659 10.1016/j.jece.2018.03.004.
660 Seiler, C. and Berendonk, T.U. (2012) Heavy metal driven co-selection of antibiotic resistance in soil
661 and water bodies impacted by agriculture and aquaculture. Frontiers in microbiology 3, 399,
662 10.3389/fmicb.2012.00399.

Sharma, V.K., Johnson, N., Cizmas, L., McDonald, T.J. and Kim, H. (2016) A review of the influence of treatment strategies on antibiotic resistant bacteria and antibiotic resistance genes. Chemosphere 150, 702-714, 10.1016/j.chemosphere.2015.12.084.

Stamm, C., Rasanen, K., Burdon, F.J., Altermatt, F., Jokela, J., Joss, A., Ackermann, M. and Eggen, R.I.L. (2016) Unravelling the Impacts of Micropollutants in Aquatic Ecosystems: Interdisciplinary Studies at the Interface of Large-Scale Ecology. Advances in ecological research 55, pp. 183-223, Academic press, 10.1016/bs.aecr.2016.07.002.

Stoob, K., Singer, H.P., Mueller, S.R., Schwarzenbach, R.P. and Stamm, C.H. (2007) Dissipation and transport of veterinary sulfonamide antibiotics after manure application to grassland in a small catchment. Environmental science & technology 41(21), 7349-7355, 10.1021/es070840e.

Storteboom, H., Arabi, M., Davis, J.G., Crimi, B. and Pruden, A. (2010) Tracking Antibiotic Resistance Genes in the South Platte River Basin Using Molecular Signatures of Urban, Agricultural, And Pristine Sources. Environmental science & technology 44(19), 7397-7404, 10.1021/es101657s.

Willems, A. (2014) The prokaryotes: Alphaproteobacteria and betaproteobacteria, pp. 777-851, Springer.

Wittmer, I.K., Bader, H.P., Scheidegger, R., Singer, H., Luck, A., Hanke, I., Carlsson, C. and Stamm, C. (2010) Significance of urban and agricultural land use for biocide and pesticide dynamics in surface waters. Water research 44(9), 2850-2862, 10.1016/j.watres.2010.01.030.

Wood, D.E., Lu, J. and Langmead, B. (2019) Improved metagenomic analysis with Kraken 2. Genome biology 20(1), 257, 10.1186/s13059-019-1891-0.

Yin, X., Jiang, X.T., Chai, B., Li, L., Yang, Y., Cole, J.R., Tiedje, J.M. and Zhang, T. (2018) ARGs-OAP v2.0 with an expanded SARG database and Hidden Markov Models for enhancement characterization and quantification of antibiotic resistance genes in environmental metagenomes. Bioinformatics 34(13), 2263-2270, 10.1093/bioinformatics/bty053.

Zhang, X.X., Zhang, T. and Fang, H.H. (2009) Antibiotic resistance genes in water environment. Applied microbiology and biotechnology 82(3), 397-414, 10.1007/s00253-008-1829-z.

690 Zurfluh, K., Hächler, H., Nüesch-Inderbinen, M. and Stephan, R. (2013) Characteristics of extended-
691 spectrum β-lactamase-and carbapenemase-producing Enterobacteriaceae isolates from rivers and
692 lakes in Switzerland. Applied and environmental microbiology 79(9), 3021-3026,
693 10.1128/AEM.00054-13.

694 **Figure and Table Legends**

695

696 **Table 1** Metagenome-assembled *bacA*-containing contigs to which taxonomy was successfully
697 assigned at genus level. Taxonomy assignment was performed using Kaiju, Kraken2, and the basic
698 local alignment tool for nucleotides (Blastn), and only contigs with consensus from all three
699 approaches at the genus level are shown. For Blastn, the quality criteria were P_{Ident} > 90.0 %, and Q_{cov}
700 > 90 %. P_{tot_bacA} indicates the proportion of *bacA* in the contig to the total *bacA* in the sample in terms
701 of reads per kilobase. P_{Ident} indicates the percentage of identical match. Q_{Cov} indicates the query
702 coverage**.**

703

704 **Figure 1. Derivation of dilution parameter (DP) from an upstream point A to the downstream**
705 **point B using the concentration of a pollute as a marker under the mass-flow assumption.**

706

707 **Figure 2. Levels (gene copies/mL) of *sul1* and *intI1* in the upstream near effluent discharge**
708 **point, and downstream river water quantified by qPCR**. Average *sul1* and *intI1* concentrations in
709 the (a) river Suze near Villeret (VIL) and (b) river Murg near Münchwilen (MUE). The dotted lines
710 are the estimated levels considering only dilution as a major driver according to the Eq.3 using
711 sodium, carbamazepine (CBZ), and 4/5-methylbenzotriazole (4/5MeBT) as conservative tracers. The
712 point of EF discharged was indicated by a light red vertical line. Symbols indicate the average and tips
713 of error bars are the lower and upper values of biological duplicates. The limit of detection for both
714 *sul1* and *intI1* is 12.5 copies/mL for all the samples shown here.

715

716 **Figure 3 Heterotrophic plate counts (CFU/mL) for clarithromycin and tetracycline multi-**
717 **resistant (CLR/TET) and sulfamethoxazole, trimethoprim, and TET multi-resistant bacteria**
718 **(SMX/TMP/TET) from the upstream and downstream river water in (a) Villeret (VIL), and (b)**
719 **in Münchwilen (MUE).** The predicted values were calculated using a selected conservative tracer
720 (i.e., sodium) according to Eq.3, and are shown as dotted lines in red, blue, or black. The limit of

721 detection (LOD) was 0.5 CFU/mL for CLR/TET multi-resistant bacteria and 5.0 CFU/mL for
722 SMX/TMP/TET multi-resistant bacteria. The LOD for SMX/TMP/TET is shown as a yellow dotted
723 horizontal line. The error bars indicate standard errors among technical triplicates.

724

725 **Figure 4** (a) Dilution parameter (*DP*) values over short downstream distance (i.e., D1 to D2 for VIL,
726 D1 to D3 for MUE) among different biological and conservative indicators; (b) ARGs and *intI1*
727 loadings at upstream near EF (US), treated wastewater effluents (EF), short downstream (D2 or 3);
728 and long downstream distance (D8) in Villeret (VIL), and (c) in Münchwilen (MUE). The treatment
729 pairs with significant difference in between were asterisked (*) in (a). The error bars represent upper
730 and lower values of biological duplicates for each gene in (b ~ c).

731

732 **Figure 5 Metagenomic analysis of effluent and river antibiotic resistomes at Villeret (VIL) and**
733 **Muenchwilen (MUE) sites for the selected sampling campaigns (VIL1, MUE2, and MUE3).** (a)
734 Venn diagram showing occurrence of antibiotic resistance gene subtypes in the treated wastewater
735 (EF) and in river water upstream (US) and downstream (D1, 0.5 km distance) of the effluent discharge
736 point and in the far downstream (D_Far, 6.8 – 13.7 km distance). The presence of ARGs were counted
737 from all three (VIL1, MUE2, MUE3) consolidated-campaigns for each treatment using the presence-
738 absence table shown in Dataset S6. (b) Shannon α-diversity of ASVs (blue) and metagenome-
739 assembled ARG subtypes (red). Procrustes analysis between ASVs (round dot symbols) and resistome
740 (blue arrow tips) where EF was included (c) and EF & D1 were excluded (d). The length of blue
741 arrows indicates the size of Procrustes errors. The error bars represent upper and lower values of
742 biological duplicates for each gene.

743

744 **Figure 6 Dynamics of prevalent and widespread metagenome-assembled ARGs along the river**
745 **continuum.** (a) The proportion of each gene among the 7 most frequently occurring and widespread
746 ARGs (*aph(3'')-I*, *aph(6)-I*, *mexT*, *tetQ*, *aadA*, *sul1*, and *bacA*). (b) and (c) Stacked bar charts of the

abundance of the 6 ARGs that were effluent-associated (omitting *bacA*); (b) relative abundance (GPM, gene per million) and (c) absolute abundance (GPL, gene per liter). Sample EF is shaded in red and the other river water samples are shaded in blue.

Figure 7 Gene arrangement on contigs containing *aph* and *sul1* genes. (a) Contigs containing *aph(3'')* and *aph(6),* retrieved from all samplings (VIL1, MUE2 & 3). (b) Contigs containing *sul1* retrieved from MUE3. All annotated genes showed > 90.0 % percent identity (P_{Ident}) at the protein level to reference proteins, using DIAMOND protein search against NCBI nr protein database. The contig IDs are italicized (e.g. *k121_XXXXXX*). P_{tot_aph} indicates the proportion of average coverage for the *aph*-containing contig to the sum of average coverages for all the *aph*-containing contigs identified in the sample. P_{tot_sul1} denotes the proportion of average coverage for the *sul1*-containing contig to the sum of average coverages for all the *sul1*-containing contigs identified in the sample. Only contigs with lengths > 1,000 bp are shown.

Table 1 Metagenome-assembled *bacA*-containing contigs to which taxonomy was successfully assigned at genus level. Taxonomy assignment was performed using Kaiju, Kraken2, and the basic local alignment tool for nucleotides (Blastn), and only contigs with consensus from all three approaches at the genus level are shown. For Blastn, the quality criteria were P_{Ident} > 90.0 %, and Q_{cov} > 90 %. P_{tot_bacA} indicates the proportion of *bacA* in the contig to the total *bacA* in the sample in terms of reads per kilobase. P_{Ident} indicates the percentage of identical match. Q_{Cov} indicates the query coverage.

Campaign	P_{tot_bacA} (%)	Contig Information			Kaiju	Kraken2	Blastn		
		Contig ID	Length (bp)	Coverage			Classification	P_{Ident} (%)	Q_{Cov} (%)
VIL1:US	2.4	k121_11403	455	4.0	*Pseudomonas* sp. Bc-h	*Pseudomonas azotoformans*	*Pseudomonas azotoformans* strain P45A	92.1	100
VIL1:US	2.2	k121_184413	333	3.6	*Pseudomonas cichorii*	*Pseudomonas cichorii* JBC1	*Pseudomonas cichorii* JBC1	97.6	100
VIL1:US	1.7	k121_761665	320	2.8	*Pseudomonas cichorii*	*Pseudomonas* spp.	*Pseudomonas cichorii* JBC1	90.2	99.1
VIL1:US	1.7	k121_867352	378	2.8	*Pseudomonas* spp.	*Pseudomonas fluorescens* SBW25	*Pseudomonas* sp. NS1(2017)	94.4	99.5
VIL1:US	3.4	k121_892755	1039	5.3	*Acidovorax temperans*	*Acidovorax* sp. 1608163	*Acidovorax* sp. 1608163	98.3	100
VIL1:US	1.5	k121_1008895	517	2.3	*Aeromonas* spp.	*Aeromonas* sp. CA23	*Aeromonas* sp. CA23	97.7	100
VIL1:EF	7.5	k121_372232	409	1.8	*Aeromonas* spp.	*Aeromonas media* WS	*Aeromonas media* strain MC64	99.5	100
VIL1:EF	20.0	k121_402216	594	4.7	*Aeromonas* spp.	*Aeromonas media* WS	*Aeromonas media* WS	99.7	100
VIL1:D1	4.9	k121_187307	4400	8.8	*Aeromonas media*	*Aeromonas media* WS	*Aeromonas media* WS	98.8	100
VIL1:D1	7.4	k121_695922	693	6.1	*Aeromonas media*	*Aeromonas media* WS	*Aeromonas media* WS	99.7	96.0
VIL1:D2	9.4	k121_274506	354	2.5	*Acidovorax* spp.	*Acidovorax* sp. 1608163	*Acidovorax* sp. 1608163	95.7	99.2
VIL1:D5	11.0	k121_307936	680	4.0	*Acidovorax* spp.	*Acidovorax* sp. 1608163	*Acidovorax* sp. 1608163	97.1	100
VIL1:D8	10.6	k121_916544	401	2.8	*Acidovorax* spp.	*Acidovorax* sp. 1608163	*Acidovorax* sp. 1608163	97.8	100
MUE2:D8	55.7	k121_6061	84229	12.0	*Aeromonas veronii*	*Aeromonas veronii* B565	*Aeromonas veronii* strain 17ISAe	93.3	95.8
MUE3:EF	69.6	k121_139490	591	3.3	*Aeromonas media*	*Aeromonas media* WS	*Aeromonas media* WS	99.5	100
MUE3:D1	11.3	k121_69935	869	3.6	*Aeromonas media*	*Aeromonas media* WS	*Aeromonas media* strain MC64	96.7	100
MUE3:D8	80.9	k121_528493	314	3.0	*Limnohabitans* sp. 63ED37-2	*Limnohabitans* sp. 63ED37-2	*Limnohabitans* sp. 63ED37-2	94.3	94.3

5. Figures and Tables

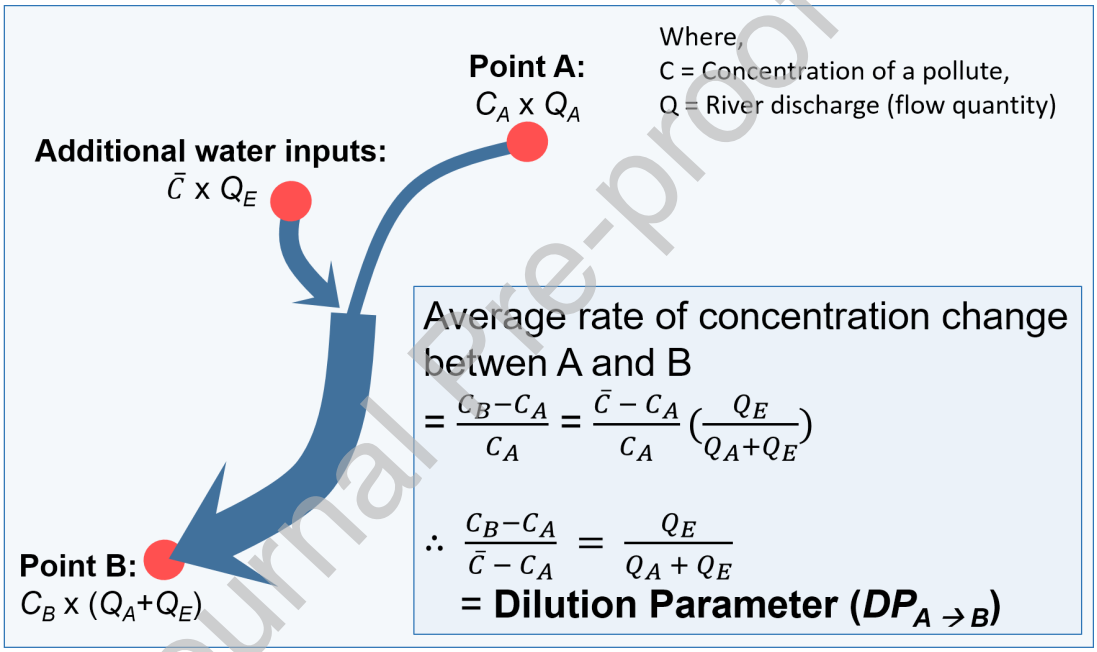


Figure 1. Derivation of dilution parameter (*DP*) from an upstream point A to the downstream point B using the concentration of a pollute as a marker under the mass-flow assumption.

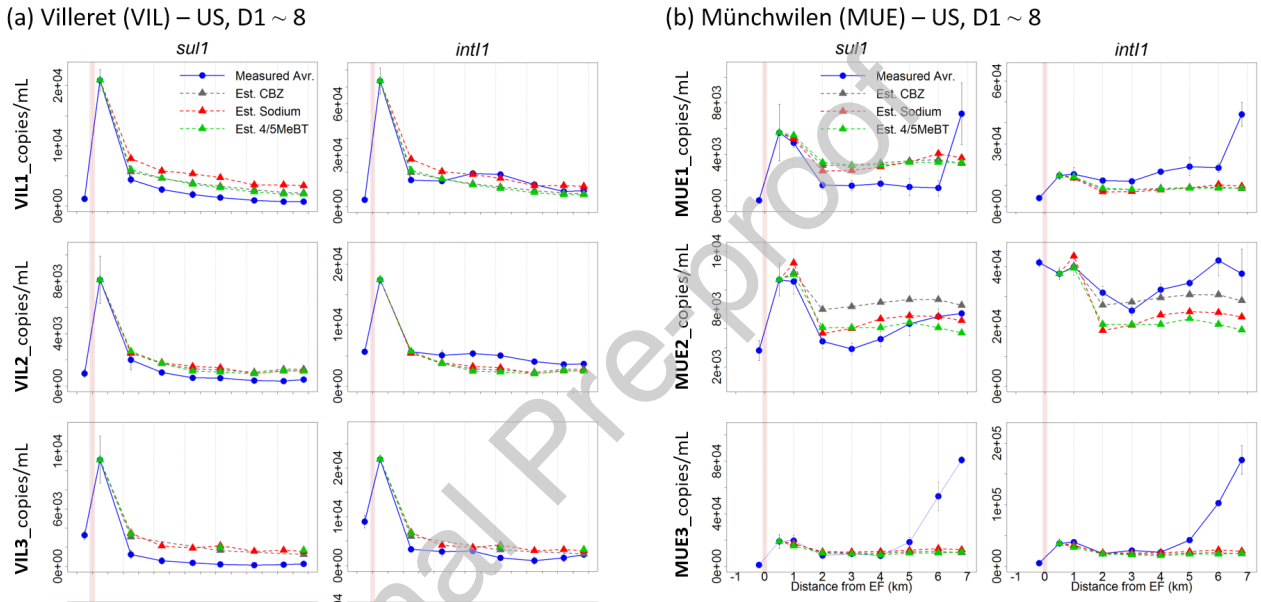


Figure 2. Levels (gene copies/mL) of *sul1* and *intI1* in the upstream near effluent discharge point, and downstream river water quantified by qPCR. Average *sul1* and *intI1* concentrations in the (a) river Suze near Villeret (VIL) and (b) river Murg near Münchwilen (MUE). The dotted lines are the estimated levels considering only dilution as a major driver according to the Eq.3 using sodium, carbamazepine (CBZ), and 4/5-methylbenzotriazole (4/5MeBT) as conservative tracers. The point of EF discharged was indicated by a light red vertical line. Symbols indicate the average and tips of error bars are the lower and upper values of biological duplicates. The limit of detection for both *sul1* and *intI1* is 12.5 copies/mL for all the samples shown here.

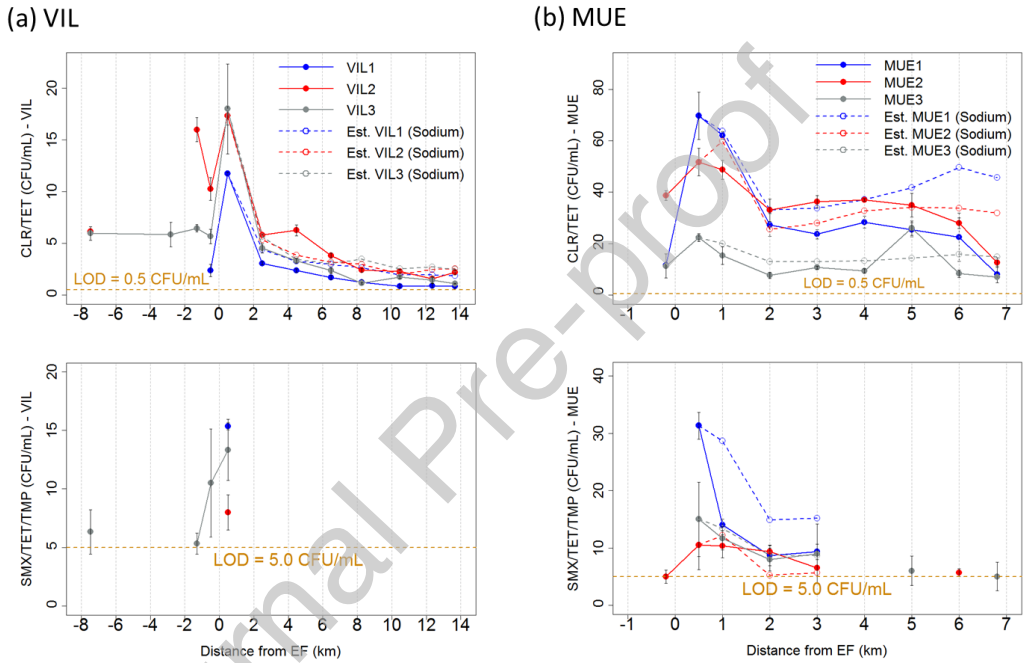


Figure 3 Heterotrophic plate counts (CFU/mL) for clarithromycin and tetracycline multi-resistant (CLR/TET) and sulfamethoxazole, trimethoprim, and TET multi-resistant bacteria (SMX/TMP/TET) from the upstream and downstream river water in (a) Villeret (VIL), and (b) in Münchwilen (MUE). The predicted values were calculated using a selected conservative tracer (i.e., sodium) according to Eq.3, and are shown as dotted lines in red, blue, or black. The limit of detection (LOD) was 0.5 CFU/mL for CLR/TET multi-resistant bacteria and 5.0 CFU/mL for SMX/TMP/TET multi-resistant bacteria. The LOD for SMX/TMP/TET is shown as a yellow dotted horizontal line. The error bars indicate standard errors among technical triplicates.

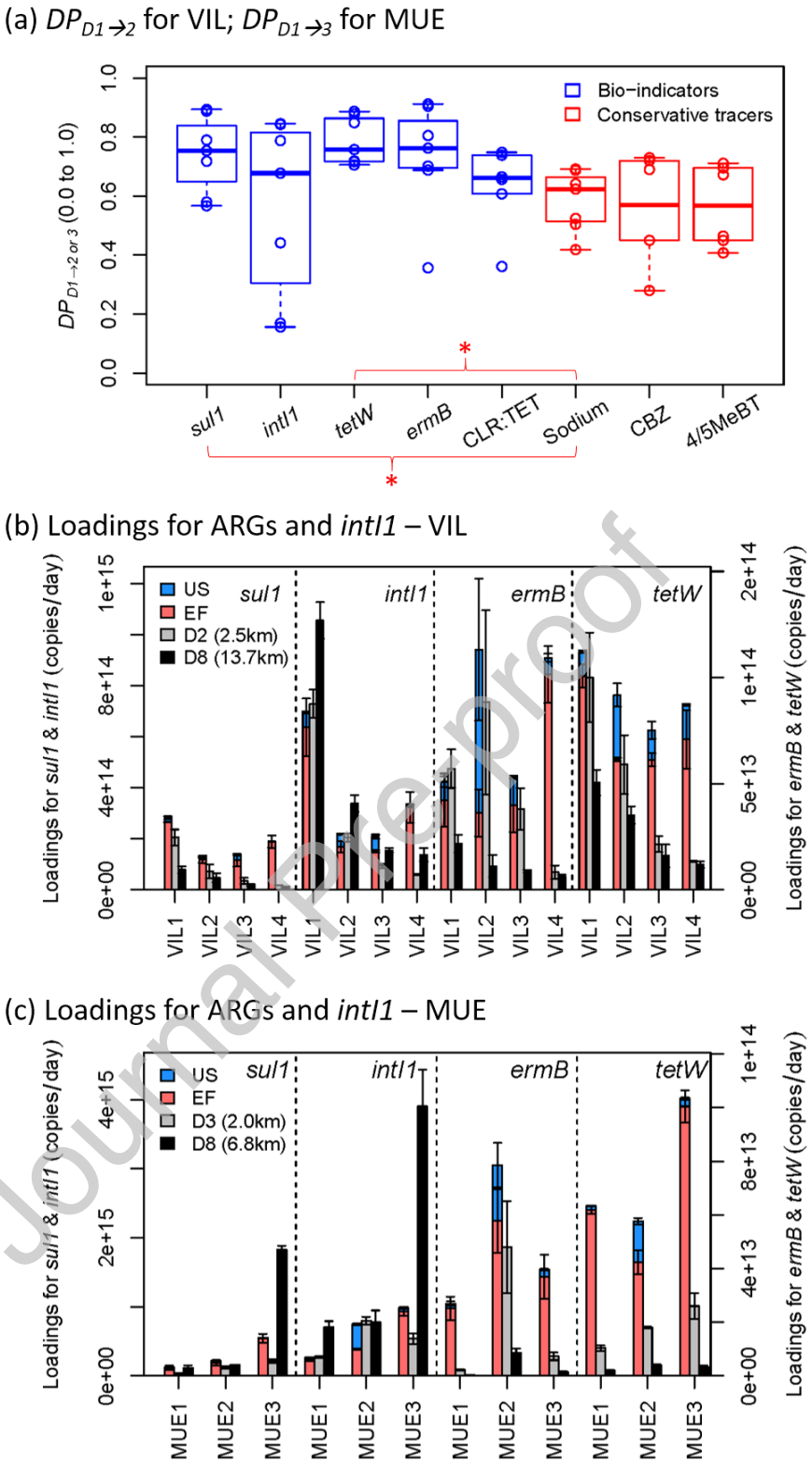


Figure 4 (a) Dilution parameter (*DP*) values over short downstream distance (i.e., D1 to D2 for VIL, D1 to D3 for MUE) among different biological and conservative indicators; (b) ARGs and *intI1* loadings at upstream near EF (US), treated wastewater effluents (EF), short downstream (D2 or 3); and long downstream distance (D8) in Villeret (VIL), and (c) in Münchwilen (MUE). The treatment pairs with significant difference in between were asterisked (*) in (a). The error bars give the upper and lower values calculated from biological duplicates for each gene in (b ~ c).

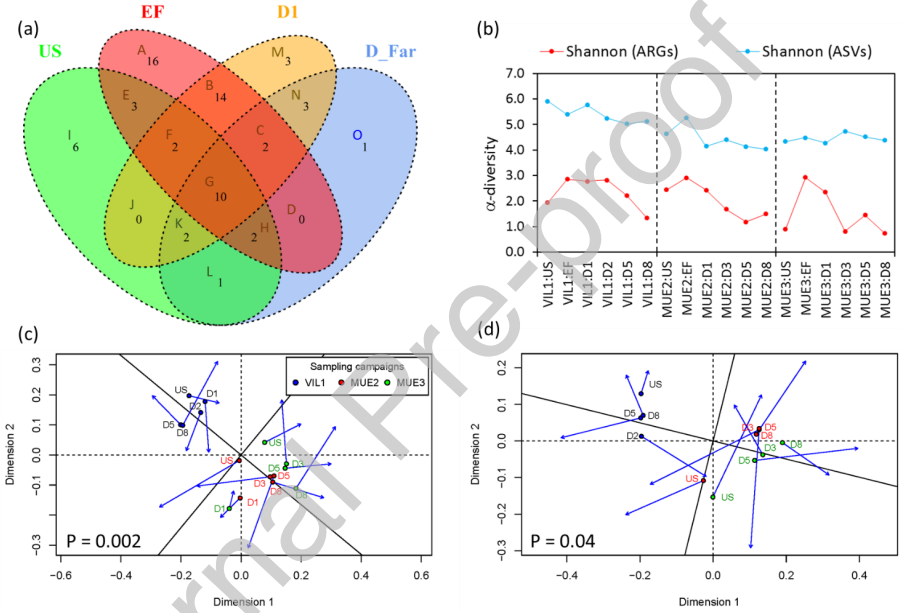


Figure 5 Metagenomic analysis of effluent and river antibiotic resistomes at Villeret (VIL) and Muenchwilen (MUE) sites for the selected sampling campaigns (VIL1, MUE2, and MUE3). (a) Venn diagram showing occurrence of antibiotic resistance gene subtypes in the treated wastewater (EF) and in river water upstream (US) and downstream (D1, 0.5 km distance) of the effluent discharge point and in the far downstream (D_Far, 6.8 – 13.7 km distance). The presence of ARGs were counted from all three (VIL1, MUE2, MUE3) consolidated-campaigns for each treatment using the presence-absence table shown in Dataset S6. (b) Shannon α-diversity of ASVs (blue) and metagenome-assembled ARG subtypes (red). Procrustes analysis between ASVs (round dot symbols) and resistome (blue arrow tips) where EF was included (c) and EF & D1 were excluded (d). The lengths of blue arrow lines indicate the size of Procrustes errors.

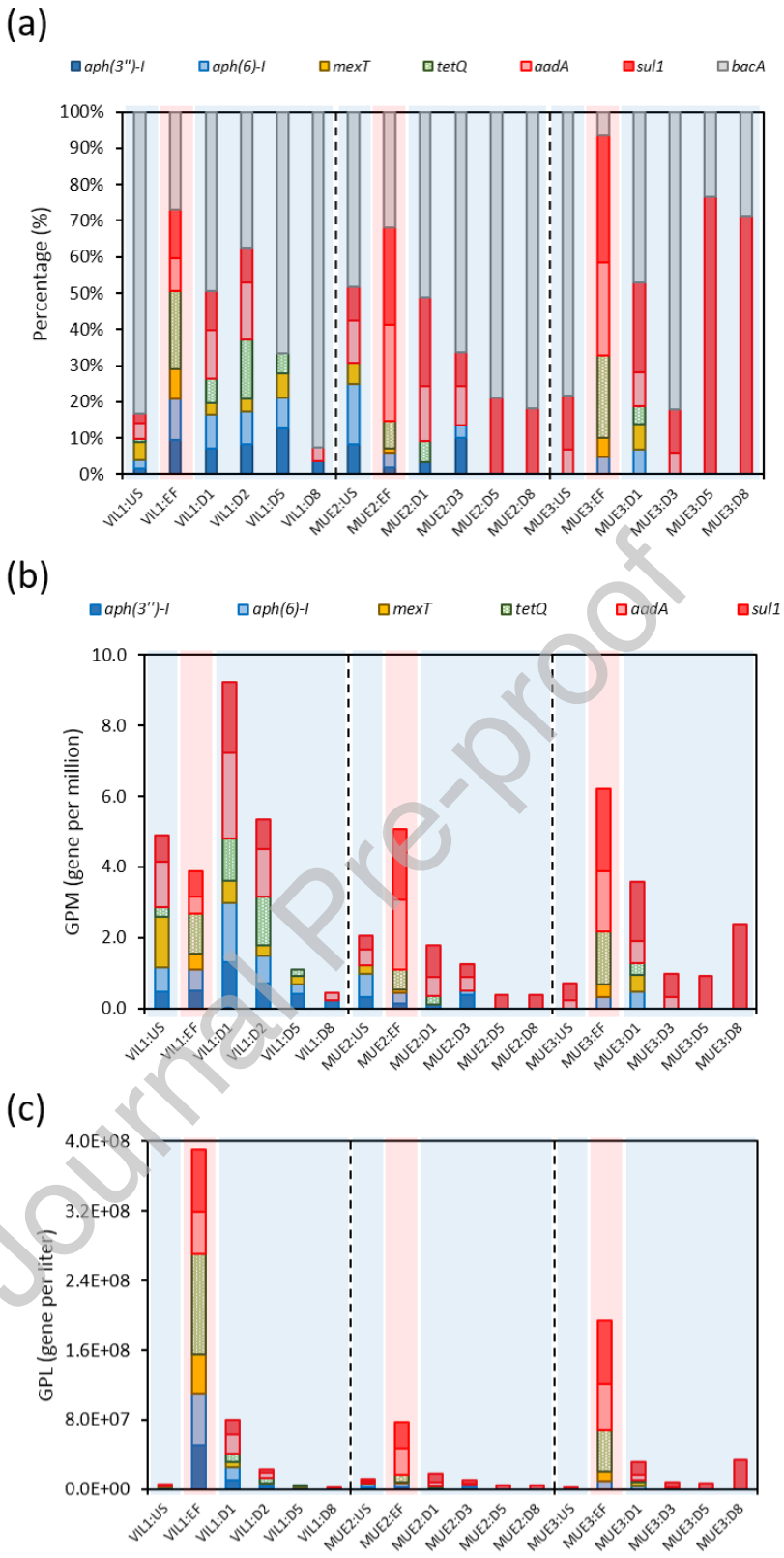


Figure 6 Dynamics of prevalent and widespread metagenome-assembled ARGs along the river continuum. (a) The proportion of each gene among the 7 most frequently occurring and widespread ARGs (*aph(3")-I*, *aph(6)-I*, *mexT*, *tetQ*, *aadA*, *sul1*, and *bacA*). (b) and (c) Stacked bar charts of the abundance of the 6 ARGs that were effluent-associated (omitting *bacA*); (b) relative abundance (GPM, gene per million) and (c) absolute abundance (GPL, gene per liter). Sample EF is shaded in red and the other river water samples are shaded in blue.

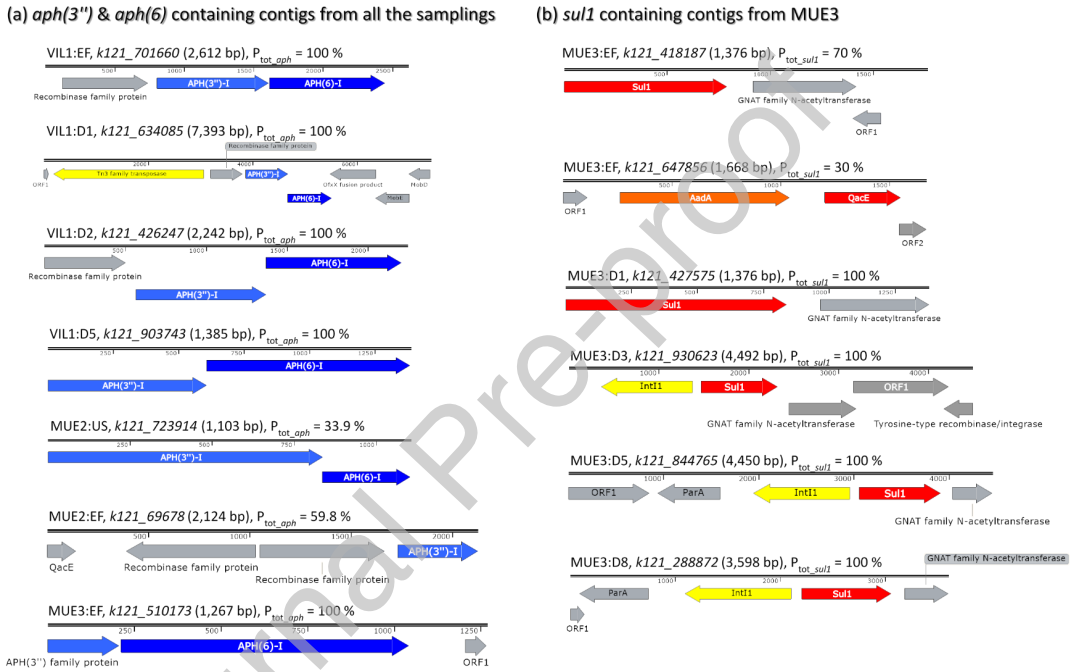
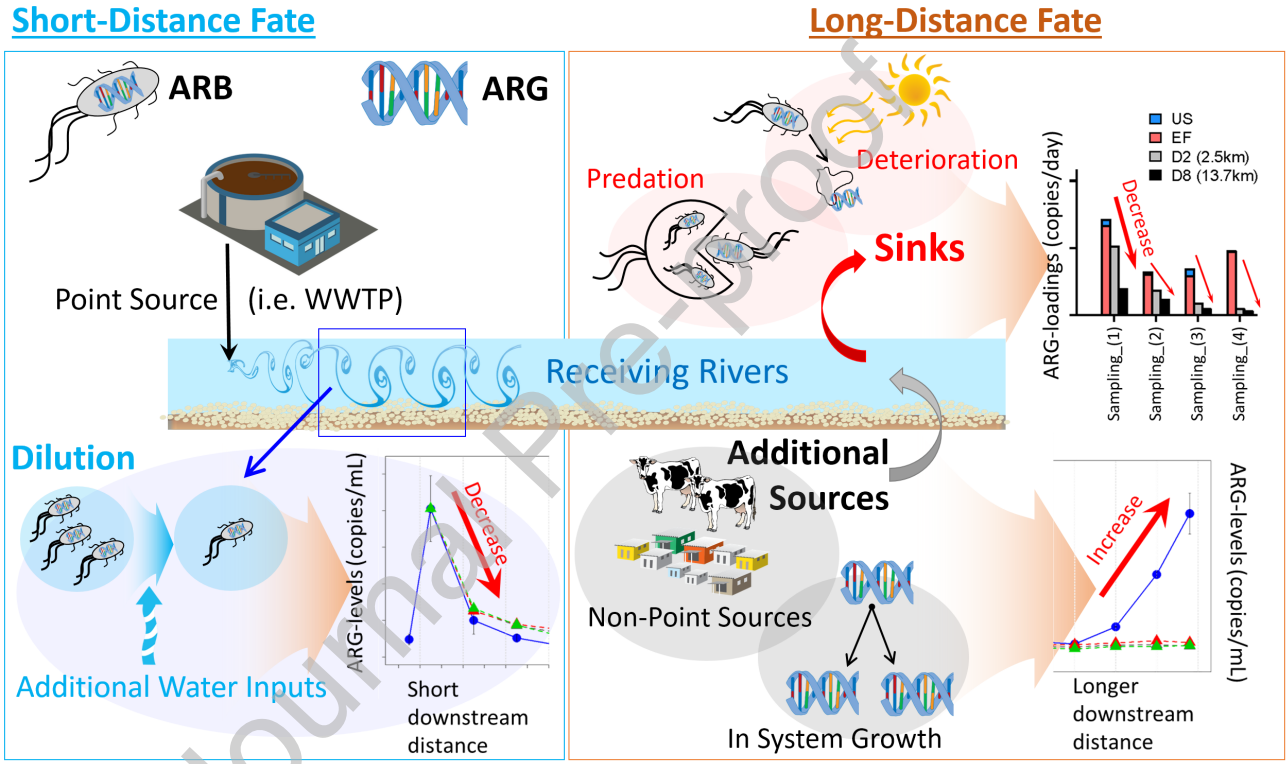


Figure 7 Gene arrangement on contigs containing *aph* and *sul1* genes. (a) Contigs containing *aph(3")* and *aph(6)*, retrieved from all samplings (VIL1, MUE2 & 3). (b) Contigs containing *sul1* retrieved from MUE3. All annotated genes showed > 90.0 % percent identity (P_{Ident}) at the protein level to reference proteins, using DIAMOND protein search against NCBI nr protein database. The contig IDs are italicized (e.g. *k121_XXXXXX*). P_{tot_aph} indicates the proportion of average coverage for the *aph*-containing contig to the sum of average coverages for all the *aph*-containing contigs identified in the sample. P_{tot_sul1} denotes the proportion of average coverage for the *sul1*-containing contig to the sum of average coverages for all the *sul1*-containing contigs identified in the sample. Only contigs with lengths > 1,000 bp are shown.

Graphical abstract



(Some) Symbols courtesy of the Integration and Application Network (ian.umces.edu/symbols/).