



**HAL**  
open science

## Abundance and environmental host range of the SXT/R391 ICEs in aquatic environmental communities

Veronica Roman, Christophe Merlin, Sandrine Baron, Emeline Larvor,  
Laetitia Le Devendec, Marko P.J. Virta, Xavier Bellanger

### ► To cite this version:

Veronica Roman, Christophe Merlin, Sandrine Baron, Emeline Larvor, Laetitia Le Devendec, et al.. Abundance and environmental host range of the SXT/R391 ICEs in aquatic environmental communities. *Environmental Pollution*, 2021, 288, pp.117673. 10.1016/j.envpol.2021.117673 . hal-03280978

**HAL Id: hal-03280978**

**<https://hal.science/hal-03280978>**

Submitted on 5 Oct 2021

**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.

1 **Abundance and environmental host range of the SXT/R391 ICEs in aquatic environmental**  
2 **communities.**

3

4 Veronica L. Roman<sup>1</sup>, Christophe Merlin<sup>1</sup>, Sandrine Baron<sup>2</sup>, Emeline Larvor<sup>2</sup>, Laetitia Le Devendec<sup>2</sup>,  
5 Marko P.J. Virta<sup>3</sup>, Xavier Bellanger<sup>1\*</sup>

6

7 <sup>1</sup> LCPME, CNRS, Université de Lorraine, Nancy, France

8 <sup>2</sup> Mycoplasmaology-Bacteriology and Antimicrobial Resistance Unit, Ploufragan-Plouzané-Niort

9 Laboratory, French Agency for Food, Environmental and Occupational Health and Safety (Anses),

10 Ploufragan, France

11 <sup>3</sup> Department of Microbiology, University of Helsinki, Helsinki, Finland

12

13

14 \* corresponding author:

15 LCPME - Equipe Microbiologie Environnementale

16 UMR 7564 CNRS - Université de Lorraine

17 Institut Jean Barriol

18 Campus Brabois Santé

19 Bâtiment AB, 3ème étage

20 9, Avenue de la Forêt de Haye - BP 20199

21 54505 Vandoeuvre-les-Nancy Cedex - FR

22 email: xavier.bellanger@univ-lorraine.fr

23

24 **Abstract**

25 Mobile genetic elements (MGEs) such as plasmids or integrative conjugative elements (ICEs) are widely  
26 involved in the horizontal transfer of antibiotic resistant genes (ARGs), but their environmental host-  
27 range and reservoirs remain poorly known, as mainly assessed through the analysis of culturable and  
28 clinical bacterial isolates. In this study, we used a gradual approach for determining the environmental  
29 abundance and host-range of ICEs belonging to the SXT/R391 family, otherwise well known to bring  
30 ARGs in *Vibrio spp.* epidemic clones and other pathogens. First, by screening a set of aquatic bacteria  
31 libraries covering 1794 strains, we found that almost 1% of the isolates hosted an SXT/R391 element,  
32 all belonging to a narrow group of non-O1/non-O139 *Vibrio cholerae*. However, when SXT/R391 ICEs  
33 were then quantified in various aquatic communities, they appeared to be ubiquitous and relatively  
34 abundant, from  $10^{-6}$  to  $10^{-3}$  ICE copies per 16S rDNA. Finally, the molecular exploration of the SXT/R391  
35 host-range in two river ecosystems impacted by anthropogenic activities, using the single-cell genomic  
36 approach epicPCR, revealed several new SXT/R391 hosts mostly in the Proteobacteria phylum. Some,  
37 such as the pathogen *Arcobacter cryaerophilus* (Campylobacteraceae), have only been encountered in  
38 discharged treated wastewaters and downstream river waters, thus revealing a likely anthropogenic  
39 origin. Others, such as the non-pathogenic bacterium *Neptunomonas acidivorans* (Oceanospirillaceae),  
40 were solely identified in rivers waters upstream and downstream the treated wastewaters discharge  
41 points and may intrinsically belong to the SXT/R391 environmental reservoir. This work points out that  
42 not only the ICEs of the SXT/R391 family are more abundant in the environment than anticipated, but  
43 also that a variety of unsuspected hosts may well represent a missing link in the environmental  
44 dissemination of MGEs from and to bacteria of anthropogenic origin.

45

46

47 **Keywords**

48 Mobile genetic elements, Aquatic environments, Resistance gene dissemination, Environmental host  
49 range

## 50 **1 Introduction**

51 The emergence and dissemination of antibiotic resistance have been recognized long ago as an  
52 unavoidable aspect of bacterial evolution following the use of antibiotics (Courvalin, 2005), which is well  
53 illustrated by the relationship between the occurrence of resistances and the antibiotic consumption  
54 (Davies, 2007; Davies and Davies, 2010; Furuya and Lowy, 2006). Mechanistically speaking, the  
55 increasing occurrence of antibiotic resistant bacteria (ARB) is mostly attributed to the selection of  
56 resistant variants that outcompete susceptible bacteria in microbial communities during antibiotic  
57 therapies (Andersson and Hughes, 2014). In parallel, antibiotic resistance genes (ARGs) also  
58 accumulate in anthropized environments, even though local antibiotic concentrations are often far below  
59 known minimal selective concentrations (Knapp *et al.*, 2010; Aubertheau *et al.*, 2017). Considering that  
60 ARB released in anthropic/animal wastes are not really fit to survive in the receiving environment, the  
61 accumulation of ARGs has been proposed to result from the horizontal gene transfer (HGT) of ARGs  
62 from released ARB to local environmental communities (Martinez, 2009). This suggests that, once  
63 released in the receiving environment, ARB from anthropic/animal wastes may survive long enough to  
64 disseminate the ARGs they carry to local environmental communities by horizontal gene transfer (HGT),  
65 reinforcing the natural ability of environmental communities to be ARGs reservoirs. Any accumulation  
66 of ARB in the environment necessarily impacts our exposure to ARGs, and increases the chances of  
67 back transfer to human/animal microbiomes, *e.g.* through vegetables grown in amended soils or  
68 recreative water, and be further selected upon new antibiotic therapies. With that respect, the  
69 dissemination of antibiotic resistances appears as an amplification loop where ARB/ARGs cycle  
70 between the environment and the animal/human microbiota, and where the increasing occurrence of  
71 ARB is fueled by each new antibiotic consumption (Xiong *et al.*, 2018). It is now widely accepted that  
72 understanding and controlling the spread of antibiotic resistances can only be achieved through a  
73 comprehensive 'One Health' approach that includes human, animal and environmental health. A more  
74 appropriate and restricted usage of antibiotics in human and animal health practices probably results  
75 from this global awareness and encouraging outcomes start to be visible (Aarestrup *et al.*, 2001).  
76 Nevertheless, over the selective effect of antibiotics, late epidemiologic studies also pointed out that the  
77 spread of ARB/ARGs seems to be the dominant factor to control (Collignon *et al.*, 2018). In that regard,  
78 identifying intermediate environmental hosts involved in the global ARG transmission pathways remains  
79 to be done. This probably lies in the fact that problematic ARGs have been isolated first because of their

80 association with pathogens. By way of consequences, much effort is given to the identification of known  
81 resistant pathogens contaminating the environment, and little case is made to the fact that, because of  
82 their association with mobile genetic elements (MGEs), *i.e.* plasmids, integrative conjugative elements  
83 (ICEs) or transposons, ARGs continue to spread toward and/or from environmental hosts by HGT. In  
84 this study, we explored the environmental microbial reservoir of ICEs belonging to the SXT/R391 family,  
85 often described as major ARG-carrying MGEs in clinical *Vibrio cholerae* strains (Spagnoletti *et al.*,  
86 2014).

87  
88 *V. cholerae* O1/O139 is the causative agent of cholera, an acute diarrheal disease, still responsible for  
89 20 000 to 140 000 deaths per year worldwide (WHO, 2020; WHO, 2019). The ICE SXT was first reported  
90 as the vector having brought ARGs in a clinical isolate of *V. cholerae* O139 isolated in India in 1993  
91 (Waldor *et al.*, 1996). Because of a conserved genetic organization (Hochhut *et al.*, 2001), SXT was  
92 grouped with the R391 element, another ICE formerly identified by Coetzee *et al.* (1972), which provides  
93 resistance genes for antibiotics and metals to a clinical strain of *Providencia rettgeri*. At present, the  
94 SXT/R391 ICEs form a family of more than 100 members sharing a core set of 43 genes, many of which  
95 being involved in DNA recombination, repair, replication and conjugative transfer (Bioteau *et al.*, 2018).  
96 In addition, SXT/R391 ICEs differ by acquisition of genetic materials either in five defined hot-spots  
97 (HS1-5) or in four variable regions (VRI-IV) of the core sequence, conferring element specific properties  
98 like metal or antibiotic resistances (Wozniak *et al.*, 2009). If the majority of known SXT/R391 ICEs  
99 integrate at the 5' end of *prfC*, a gene coding for the peptide chain release factor 3 (Hochhut and Waldor,  
100 1999), a recent study demonstrated that some others integrate at the 3' end of a *tRNA-Ser* encoding  
101 gene (Bioteau *et al.*, 2018). Hence, the site-specific integration into the 5' end of *prfC* is no more a  
102 standard to be part of this family, which is now divided in 4 groups according to their insertion sites,  
103 together with their structural organization and their distribution in bacterial species (Bioteau *et al.*, 2018).

104  
105 To date, SXT/R391 ICEs were mostly found in *Vibrio* and *Proteus* species from human/animal bacteria  
106 isolated based on their resistance profile and molecular typing (Lei *et al.*, 2016; Sato *et al.*, 2020). These  
107 ICEs have not been widely searched for in environmental species, although they have been identified  
108 in a few environmental gammaproteobacteria like *Shewanella*, *Marinomonas* or *Actinobacillus*  
109 (Pembroke and Piterina, 2006; Badhai and Das, 2016; Li *et al.*, 2018). To the best of our knowledge,

110 SXT/R391 ICEs have been mainly (i) detected through culture-dependent approaches, (ii) from clinical  
111 isolates, and (iii) using molecular typing targeting the integrase gene of the element, which seems now  
112 to be obsolete (Balado *et al.*, 2013; Lei *et al.*, 2016; Shinde *et al.*, 2020; Sato *et al.*, 2020). As a matter  
113 of fact, our knowledge regarding the ecology of SXT/R391 ICEs is very limited and the prevalence of  
114 these elements in the environment is mostly unknown.

115

116 In this study, we explored the abundance and host range of SXT/R391 ICEs in environmental aquatic  
117 communities and wastewaters using culture-based and molecular-based approaches in a gradual  
118 analysis. First, four collections of bacterial strains isolated in France from environmental waters,  
119 wastewaters and fishes were screened for the presence of SXT/R391 ICEs using new PCR  
120 primer/probe designs. The environmental abundance of SXT/R391 ICEs was then determined by  
121 quantitative real time PCR (qPCR) for identifying possible reservoirs in some French rivers, wastewaters  
122 and fish associated communities. Finally, the environmental hosts of the SXT/R391 ICEs were further  
123 identified using a state-of-the-art technique referred as epicPCR (Emulsion, Paired Isolation and  
124 Concatenation PCR). This approach revealed that SXT/R391 ICEs are quite common in environmental  
125 aquatic communities and wastewaters, and allowed us to identify new bacterial host species of this  
126 family of elements that could well constitute environmental reservoirs of these ICEs.

127

128

## 129 **2. Materials and Methods**

### 130 **2.1 Sample collection and bacteria libraries**

131 Water and fish samples were first collected from two fish farming facilities differing by their fish cultures  
132 and their water circulation systems. The first one, managed by the DAC team of the UR AFPA laboratory  
133 based in Nancy (North-East of France), grows sturgeon, perch and zander in controlled systems fed by  
134 the same tap-water source that is independently treated and re-circulated in each pond. The second  
135 one, managed by the PVP from the Anses laboratory based in Ploufragan-Plouzané (North-West of  
136 France), grows rainbow-trout in semi-open raceways fed by natural fresh water. These two sample sets  
137 were collected on July 2019 and December 2018, respectively. Water samples were also taken from  
138 two different municipal treatment plants in Nancy and Joeuf cities, located in the North-East of France,  
139 and the Meurthe and the Orne Rivers which are their respective receiving rivers. These samples were

140 collected on November 2017 and September 2019 in the Meurthe River ecosystem, and on September  
141 2018, March 2019 and September 2019 in the Orne River ecosystem. The wastewater treatment plant  
142 (WWTP) located in Nancy processes around 12 000 m<sup>3</sup> of wastewater a day from 120 000 residents,  
143 while the one located in Joeuf processes around 7 700 m<sup>3</sup> of wastewater a day from 30 000 residents.  
144 Water was also collected from Mont-sur-Meurthe (MSM), Damelevières (DAM), Saint-Nicolas-de-Port  
145 (SNP), Art-sur-Meurthe (ASM) and Nancy (NAN) which are cities located along a 30 km long section of  
146 the Meurthe River, and from Joeuf (Upstream WWTP) and Beth barrage (Downstram WWTP) on the  
147 Orne River. Finally, this study also focused on the SXT/R391 content of 1794 bacteria isolated from four  
148 aquatic environments and wastewaters at the Anses Ploufragan-Plouzané laboratory. The 4  
149 corresponding strain libraries mainly containing *V. cholerae* non-O1 non-O139, *Aeromonas spp.*, and  
150 *Pseudomonas spp.* strains collected over 19 years in France (Table S1). *Vibrio cholerae* and  
151 *Aeromonas spp.* strains were isolated and identified as described in Baron *et al.* (2013) and Baron *et*  
152 *al.* (2017), respectively. *Pseudomonas spp.* strains were isolated on RAPID'P.aeruginosa Agar (Bio-  
153 Rad) before being identified by PCR according to Benami *et al.* (2013). For epicPCR, 200 ml of water  
154 samples were centrifuged for 40 min at 8,000 g and 4°C. After discarding the supernatant, the remaining  
155 cell pellets were resuspended in 20% glycerol and stored at -20°C until use.

156

## 157 **2.2 DNA extraction**

158 For total community DNA extractions, 100 mL of water samples, or *ca.* 100 mg of fish samples were  
159 dispersed in 100 mL of apyrogenic water (BBraun), were filtered on polycarbonate membranes  
160 (Whatman Nucleopore filter, pore size 0.22 µm, diameter 47 mm) using a filtration device (Combisart 6-  
161 branch Manifold, Sartorius) and stored at - 20°C until DNA extraction. Total community DNAs from  
162 stored filters were extracted using the "PowerWater DNA Isolation Kit" (MO BIO laboratories Inc)  
163 according to the manufacturer recommendations. DNAs were eluted from silica columns with 100 µL of  
164 PCR grade water (RNase-Free Water, Qiagen). Genomic DNAs from bacterial isolates (libraries) were  
165 isolated using the InstaGene Matrix (BioRad) according to the manufacturer recommendations. DNA  
166 concentrations and purities were estimated by spectrophotometry according to standard procedures  
167 and/or using a QuBit fluorometer.

168

## 169 **2.3 Primers design and specificity checking**

170 The ICE SXT<sub>MO10</sub> from *V. cholerae* O139 MO10 (GenBank: AY055428.1) (Beaber *et al.*, 2002b) was  
171 used as reference in order to browse databases and recover sequenced SXT/R391 ICEs using BLASTn  
172 (Altschul *et al.*, 1997) in September 2017. All collected sequences were confirmed to be part of  
173 SXT/R391 elements by the presence of *prfC*, *int<sub>prfC</sub>* and/or *setR* genes, corresponding to the ICE *attB*  
174 site, the ICE integrase gene and the ICE master regulatory gene, respectively. The SXT/R391  
175 sequences corresponding to the conserved *traB* and *int<sub>prfC</sub>* genes were aligned using ClustalW and, for  
176 each, a couple of primers and a Taqman probe were designed (Table S2). Primer specificities were  
177 checked thanks to endpoint PCRs targeting SXT/R391 *traB* or *int<sub>prfC</sub>* genes on total community DNA  
178 extracted from the Meurthe River water. These PCRs were performed using Phusion polymerase  
179 (Thermo Fisher Scientific) with the following cycling conditions: initial denaturation at 98°C for 30s, then  
180 35 cycles at 98°C for 10 s, 60°C for 30 s and 73°C for 15 s, and a final extension for 10 min at 72°C.  
181 The amplicons were cloned using the “Zero Blunt TOPO PCR Cloning Kit” (Invitrogen) before the inserts  
182 of 100 recombinant plasmids were sequenced (Genewiz) for confirming their SXT/R391 origin.

183

#### 184 **2.4 Quantitative PCR assays**

185 SXT/R391 elements were detected and quantified by qPCR by targeting either the *traB* or the *int<sub>prfC</sub>*  
186 genes from genomic and total environmental DNAs. The *traB* gene has been preferentially used as  
187 SXT/R391 ICEs exhibiting different *int* gene variants have been reported (Bioteau *et al.*, 2018).  
188 Quantitative PCRs were performed in triplicate using a “StepOnePlus Real-Time PCR System” (Applied  
189 Biosystems) with the “Taqman Universal Master Mix No UNG” when targeting the SXT/R391 ICEs, or  
190 with the “Power SYBR® Green PCR Master Mix” (Applied Biosystems) when the targeting class 1  
191 integrons and 16S rRNA genes. SXT/R391 elements were quantified by qPCR according to the kit  
192 manufacturer, using primers at 40 µM (Table S2) and the following cycling conditions: initial denaturation  
193 at 50°C for 2 min, followed by 45 cycles at 95°C for 15s and 60°C for 1 min for annealing and extension.  
194 Class 1 integrons and eubacterial 16s rRNA genes were quantified according to De la Cruz Barrón *et al.*  
195 *et al.* (2018). Plasmids pBELX2, containing a *traB* gene fragment from ICE SXT<sub>MO10</sub>, or pNORM1,  
196 containing *intI1* and 16S rRNA gene fragments (Rocha *et al.*, 2020), linearized with either *NdeI* or  
197 *BamHI*, were used as qPCR standards for the quantification of SXT/R391 ICEs or class 1 integrons and  
198 environmental bacteria, respectively. These plasmids were extracted using the “Wizard plus SV  
199 Minipreps DNA Purification System” (Promega) according to the manufacturer recommendations. Their



200 DNA concentrations and purities were estimated by spectrophotometry according to standard  
201 procedures and/or using a QuBit fluorometer. The absence of residual inhibitors in the DNA extracts  
202 was checked by comparing qPCR amplifications from serially diluted DNA templates.

203

#### 204 **2.5 16S rRNA gene metagenomics**

205 For the Orne River water samples, the PCR amplifications and sequencing of the V3-V4 region of the  
206 16S rRNA genes were performed on total microbial DNAs at the Institute of Biotechnology, University  
207 of Helsinki, Finland, according to Hultman *et al.* (2018). For the Meurthe River water samples, PCR  
208 amplification of the V3-V4 region of the 16S rRNA genes on total microbial DNA was done with the 515F  
209 and 806R primers (Table S2) using Phusion polymerase and HF buffer (New England Biolabs). The  
210 cycling conditions were as follows: initial denaturation at 98°C, followed by 28 cycles at 98°C for 10 s,  
211 58°C for 30 s and 72°C for 15 s, and a final extension for 5 min at 72°C. The PCR products were pooled,  
212 cleaned (Monarch PCR & DNA Cleanup Kit) and sequenced on Illumina Miseq (2x250) (Genewiz). All  
213 16S rRNA gene sequences have been deposited in GenBank (BioProject accession no. SUB8106542  
214 and SUB8106791).

215

#### 216 **2.6 EpicPCR**

217 EpicPCR (Emulsion, Paired Isolation and Concatenation PCR) is a recent single-cell genomic method,  
218 based on fusion-PCR amplifications performed on bacteria isolated in polyacrylamide beads. This  
219 method allows linking a functional sequence of interest to a 16S rRNA gene fragment, before  
220 sequencing for the taxonomic assignment of the bacteria carrying the functional sequence. Before the  
221 polyacrylamide beads formation step, cells were stained with SYBR Green I (Invitrogen), filtered on a  
222 track-etched polycarbonate membrane (Whatman Nucleopore, 0,2 µM) and enumerated on an  
223 epifluorescence microscope. The epicPCR amplifications were done as described by Hultman *et al.*  
224 (2018) with the modifications proposed by Roman *et al.* (2021). The PCR products were pooled, purified  
225 with the Monarch PCR & DNA Cleanup Kit (NEB) following the manufacturer recommendations.  
226 Thereafter, they were either directly sequenced on the Illumina Miseq (2x250) platform (Genewiz) or  
227 cloned using the Zero Blunt TOPO PCR Cloning Kit (Invitrogen) as previously described, before  
228 approximately 100 clones were sent for Sanger sequencing.

229

## 230 **2.7 Sequencing data analysis**

231 For recombinant plasmids, inserts were sequenced using the Sanger method (Genewiz) using primers  
232 listed in Table S2. The sequences obtained were analyzed using BLASTn (Altschul *et al.*, 1997). For  
233 metagenomics, 16S rDNA reads were joined with Pear (Zhang *et al.*, 2014) using default options,  
234 sequence quality was checked with FastQC (Andrews, 2020), primer sequences and short reads (< 330  
235 bp) were removed using cutadapt (Martin, 2011). Unique sequences were identified using USEARCH -  
236 fastx\_uniques command (Edgar, 2010). OTUs were clustered and reads were mapped to reference  
237 sequences with the USEARCH -cluster\_otus command with the -minsize 2 parameter. Taxonomic  
238 classification of OTUs was done using the USEARCH -sintax command, using the ltp\_16s\_v123  
239 database with the classifier cutoff set at 60. A representative sequence from each bacterial family  
240 (constituted of at least 10 reads) was aligned with MUSCLE (Madeire *et al.*, 2019), before a phylogenetic  
241 tree was built with Fasttree (Price *et al.*, 2009) and visualized in iTol (Letunic and Bork, 2019). Reads  
242 resulting from epicPCR amplifications were analyzed as described by Roman *et al.* (2021) and a  
243 confidence level (CL) in the detection of OTU determined, from level 1 (the lower), to level 5. This  
244 corresponds to the validation of an OTU based on its detection in epicPCR replicates as follows: (i) in  
245 only one replicate (CL=1), (ii) in either 2 out of 2 replicates of technical duplicates (so called a “2/2  
246 technical replicates” validation method) or “2/3 technical” replicates (CL=2), (iii) in “3/3 technical”  
247 replicates (CL=3), (iv) in either “2/2 biological” or “2/3 biological” replicates (CL=4) or (v) “3/3 biological”  
248 replicates (CL=5).

249

250

## 251 **3 Results and Discussion**

### 252 **3.1 Design of new set of SXT/R391 targeting primers**

253 Until now, the vast majority of PCR-based identifications of SXT/R391 ICEs were performed by targeting  
254 the *int<sub>prfC</sub>* gene coding for the integrase of elements integrated in the *prfC* gene 5' end (Balado *et al.*,  
255 2013; Lei *et al.*, 2016; Shinde *et al.*, 2020). Only recently, some authors focused on the conserved *rumA*  
256 gene (Sato *et al.*, 2020). Another recent study highlighted that SXT/R391 elements can carry an  
257 alternative integrase gene, signifying that *int<sub>prfC</sub>* is no more considered as part of the core sequence of  
258 the entire SXT/R391 ICEs (Bioteau *et al.*, 2018). Hence, a new SXT/R391 PCR target was here

259 searched for in order to broadly and specifically detect/quantify these ICEs in isolates and environmental  
260 samples, and even further identify new environmental hosts by epicPCR. In this context, we designed  
261 a new set of primers/probe targeting the *traB* gene, itself part of the SXT/R391 core sequence, coding  
262 for a protein involved in the conjugation process, and only presents in one copy per ICE (Table S2).  
263 Browsing database allowed verifying that the *traB* targeted fragment was only found in members of this  
264 ICE family, despite the fact that more globally, the *tra* genes of SXT/R391 ICEs are related to those of  
265 the IncF plasmids (Beaber *et al.*, 2002a). Further verifications of the primer design were performed by  
266 amplifying, cloning and sequencing amplicons obtained by PCR performed on total genomic DNA of  
267 microbial communities sampled in the Meurthe River. The 34 sequences obtained were identified by  
268 multi-alignment analysis as corresponding to 5 different *traB* sequence variants. The three most  
269 abundant of these variants were so far not listed in databases (Table S3). In spite of this novelty, the 5  
270 sequence variants were confirmed to actually belong to the *traB* gene of SXT/R391 ICEs. Indeed, the  
271 BLASTn analyses of each of 5 sequence variants with the parameters "query cover" > 99% and "percent  
272 identity" > 90% allow identifying 206 sequences. These sequences include 205 ICE fragments  
273 annotated as such or chromosome fragments, therefore likely ICEs, and a single IncF plasmid fragment  
274 belonging to plasmid pRHBSTW-00054\_2 and identical to sequence variant n°2 also detected in  
275 ICEVchCHN4210 (Table S3). Nevertheless, other IncF plasmids were not identified with the parameters  
276 used, since they have both query cover and percent identity  $\leq$  83%. Given the relationship between the  
277 conjugation modules of SXT/R391 ICEs and IncF plasmids, it is likely that, with the IncF plasmid  
278 pRHBSTW-00054\_2, we identified an allelic exchange between a part of the plasmid conjugation  
279 module and the one of an SXT/R391 ICE closely related to ICEVchCHN4210. Whatever, the isolation  
280 of 3 *traB* new sequences variant suggests a broader diversity of SXT/R391 ICEs sequence in the  
281 environment than up to now reported. Finally, qPCR amplifications performed on environmental DNA  
282 with, either a SYBR Green or a Taqman chemistry, resulted in the detection of a single and expected  
283 81bp-band in an agarose gel electrophoresis (Figure S1), which further validate the proper specificity of  
284 the designed primers/probe.

285

286

287 **3.2 Abundance of STX/R391 ICEs in environmental bacteria libraries**

288 SXT/R391 elements were first discovered in a clinical strain of *V. cholerae* O139 (SXT<sub>MO10</sub>) (Beaber *et*  
289 *al.*, 2002b) and thereafter most of the time reported in other *Vibrio* strains/species (Bioteau *et al.*, 2018).  
290 We therefore first screened two libraries (Table S1, libraries #2 and #3) made after the sampling of  
291 aquatic environments supporting these bacteria and having indeed proved to contain 168 non-O1/non-  
292 O139 *V. cholerae* strains. These environments were fresh waters sampled in the Rance and the  
293 Queffleuth Rivers (North-Western France), raw and treated wastewaters, marine water, and brackish  
294 water (see Table S1 for details). The detection of SXT/R391 elements was performed using a qPCR  
295 targeting the conserved *traB* gene and resulted in spotting 14 SXT/R391 carrying strains, which  
296 represent 8% (14/168) of the total *V. cholerae* collection, all recovered from the Rance River  
297 environment. Thirteen of the 14 detected SXT/R391 ICEs were further characterized with a qPCR  
298 targeting the *int<sub>prc</sub>* integrase gene. We determined that 8 of these ICEs carried the *int<sub>prc</sub>* gene (Table  
299 S4), suggesting that the 5 remaining strains may host an ICE integrated in an alternative *attB* site, such  
300 as the already identified *tRNA-Ser* encoding gene (Bioteau *et al.*, 2018). This again reinforces the idea  
301 that ICEs from the SXT/R391 family are numerous and diverse.

302 The literature reported that non-*Vibrio* SXT/R391 carriers, all belonging to gamma-proteobacteria,  
303 including *Proteus spp.* or *Shewanella spp.* (Bioteau *et al.*, 2018), were sporadically isolated from aquatic  
304 environments, stool samples and infected patients, and that some are known to be human- or animal-  
305 pathogens, notably fishes. For this reason, we also screened other libraries containing 1626 strains from  
306 42 genera (of which 1436 gamma-proteobacteria but no *Vibrio sp.*), isolated from environmental waters,  
307 wastewaters and wild fishes, but none of them was shown to carry any SXT/R391 element. At this stage,  
308 it is worth mentioning that, among the 1626 characterized strains, 443 were *Aeromonas spp.* isolates.  
309 All in all, if the new gene target *traB* allowed detecting SXT/R391 elements in a few environmental non-  
310 O1/non-O139 *V. cholerae* strains, none were identified in strains belonging to any other species/genus,  
311 which tends to comfort data from literature suggesting that SXT/R391 elements are mainly hosted by *V.*  
312 *cholerae* and other *Vibrio spp.* strains. Nevertheless, it is well known that culture-based approaches  
313 only capture a partial picture of the reality. To go further, we decided to investigate the environmental  
314 host range of the SXT/R391 elements using epicPCR. To do so, it was first necessary to assess the  
315 feasibility of the approach by determining the relative abundance of the SXT/R391 ICE family in the  
316 environment. In addition, such abundance study would valuably provide insight on the type of

317 environments permissive to the SXT/R391 family, as well as a better vision of the significance of the  
318 SXT/R391 ICEs in the environment.

319

### 320 **3.3 Abundances of SXT/R391 ICEs in aquatic environments**

321 To date, all known hosts of SXT/R391 ICEs are either waterborne bacteria or human/animal intestinal  
322 bacteria that may entered aquatic environments through the discharge of WWTP effluents or agricultural  
323 activities such manure/slurry spreading. In this context, we investigated the relative abundances of  
324 SXT/R391 ICEs in various aquatic ecosystems impacted by Human activities. We first studied the  
325 abundance of these MGEs in two fish farming facilities differing by their fish cultures and water  
326 circulation systems. Water and biofilm samples were collected from sturgeon, perch, zander and  
327 rainbow-trout growing basins. The stomach contents, intestine contents and feces of three rainbow-trout  
328 were also collected (details in Supplementary material). Total community DNAs were extracted from  
329 these samples, and their contents in SXT/R391 ICEs were determined by qPCR targeting the *traB* gene  
330 as described above. The SXT/R391 ICE quantities in communities were further normalized by the 16S  
331 rDNA abundance, also estimated by qPCR, so as to get a comparable assessment of the ICE  
332 abundance. The relative abundances of SXT/R391 ICEs in water, biofilm and fish samples are  
333 presented in Figure 1A. SXT/R391 ICEs were detected in all analyzed communities and their relative  
334 abundances range from  $10^{-6}$  copies per 16S rDNA in fishpond-associated biofilm, to almost  $10^{-3}$  copies  
335 per 16S rDNA in fish feces. Concerning the recirculation systems, no significant variation of SXT/R391  
336 ICEs relative abundances was observed after the mechanical filtering nor the UV-treatment steps,  
337 indicating that those steps do not have any visible effect on the abundance of SXT/R391 ICEs. The  
338 comparison of SXT/R391 ICE abundance in the different basins showed almost 10 time more ICEs in  
339 water samples from Perch basins, suggesting a possible host-fish effect on the abundance of SXT/R391  
340 elements. In the raceway systems containing rainbow-trout, the relative abundance of SXT/R391 ICEs  
341 is from 1 to 2 orders of magnitude lower in water samples than in fish samples, indicating that these  
342 animals may act as reservoirs and vectors of SXT/R391 carrying bacteria, from where they would further  
343 disseminate in their surrounding environment.

344 Taking these observations into account, we then studied the abundance of SXT/R391 ICEs in water  
345 samples from two rivers impacted by the discharge of urban WWTP effluents. The relative abundance  
346 of SXT/R391 ICEs in water sample sets collected upstream and downstream two WWTPs located on

347 the bank of the Orne and Meurthe Rivers as well as in the corresponding effluents are presented Figure  
348 1B. Given that there is no increase of the relative abundance of SXT/391 ICEs after the WWTPs in both  
349 rivers, we conclude that, in these ecosystems, there was no observable effect of the WWTP effluent  
350 discharge on the global SXT/391 content of the receiving river, even for the first sampling day on the  
351 Orne River when abundance river water in upstream WWTP and in its effluents were very different. This  
352 absence of quantitative effect is likely due to the high dilution of the WWTP effluents in the studied river  
353 waters, since, in this area of France, there is no real water scarcity. However, this apparent quantitative  
354 stability does not presume of a substantial change in the river microbial community composition of the  
355 river water after the WWTP discharge points.

356 The Meurthe River flows close to industrial facilities exploiting rock salt and discharging salted residual  
357 water in the river, therefore creating a salt concentration gradient along the course of the river, which  
358 may favor community changes. The SXT/R391 elements were quantified at five different locations along  
359 a 30 km-long section of the Meurthe River encompassing the Saint-Nicolas-de-Port and Art-sur-Meurthe  
360 urban areas where the rock salt exploiting industries are located. The relative abundances of class 1  
361 integrons, usually used as a proxy for anthropogenic pollution, were also determined in order to get a  
362 reference with a broadly used MGE proxy. The relative abundance of SXT/R391 ICEs and class 1  
363 integrons in the Meurthe River are presented Figure 2, where the corresponding conductivity of the  
364 water samples provides an idea of the evolution of the salt concentration. As expected, there is an  
365 increase in the river water conductivity values starting from Saint-Nicolas-de-Port (SNP) to Nancy  
366 (NAN). The relative abundances of class 1 integrons in those water samples are 13 to 140 times higher  
367 than the ones of SXT/R391 ICEs. But, the relative abundance of SXT/R391 elements gradually increase  
368 by approximately 1 order of magnitude from upstream to downstream, while the relative abundance of  
369 class 1 integrons remained stable. Statistical analyses allowed pointing out that (i) the relative  
370 abundances of SXT/391 elements evolved independently from those of class 1 integrons (Spearman's  
371 rank correlation test:  $P=0.78$ ), (ii) the relative abundance of class 1 integrons does not correlate water  
372 conductivity (Spearman's rank correlation test:  $P=0.78$ ), while (iii) the relative abundance of SXT/391  
373 elements does (Spearman's rank correlation test:  $P=0.02$ ), thus highlighting salinity as an important  
374 driving parameter for SXT/R391 hosts ecology.

375 Regarding the data obtained on the sole Orne River (Figure 1B), the results obtained at two different  
376 sampling dates showed that these abundances varied up to two orders for a same sampling site,

377 suggesting a sampling date effect (ANOVA test,  $P=0.12$ ). When considering the data obtained from the  
378 different sampling campaigns performed on the Meurthe River (Figures 1 and 2), we were not able to  
379 detect any sampling date effect (Student T test for paired values,  $P=0.26$ ). Anyway, it can be noticed  
380 that SXT/R391 elements tend to be more abundant in the Meurthe River (from  $4.4 \times 10^{-5}$  to  $7.3 \times 10^{-4}$   
381 copies/16S rDNA) than in the Orne River (from  $1.9 \times 10^{-7}$  to  $8.7 \times 10^{-5}$  copies/16S rDNA). The latter result  
382 raises the question of the higher abundance of SXT/R391 ICEs in the Meurthe River, as well as its  
383 increase along the course of the river together with salinity. This enrichment could well relate to  
384 community shifts originating from changes in environmental conditions favoring some meso/polyhaline  
385 bacterial species hosting SXT/R391 elements, such as *Vibrio cholerae*, as oppose to oligohaline  
386 bacteria. Other meso/polyhaline bacteria species that do not belong to the *Vibrio* genus may also be  
387 involved in this phenomenon.

388 Finally, as they have been detected in all tested environmental samples, SXT/R391 ICEs appeared to  
389 be widespread and likely ubiquitous. A very recent paper also demonstrate that the elements of this  
390 family are likely very abundant in WWTP effluents (Calderón-Franco *et al.*, 2020), even suggesting they  
391 would be one of the most abundant ICE family.

392

### 393 **3.4 Identification of SXT/R391 hosts in the Orne and Meurthe River ecosystems by epicPCR**

394 Having detected the SXT/R391 elements in large enough abundances in the Orne and Meurthe Rivers,  
395 as well as in WWTP effluents, their host range was further investigated using the cutting-edge single  
396 cell metagenomic approach epicPCR (Emulsion, Paired, Isolation and Concatenation PCR). By doing  
397 so, we attempted to determine both if SXT/R391 hosts from anthropogenic sources could be detected  
398 in rivers downstream WWTP effluent discharge point, despite the fact that the global abundance of these  
399 ICEs remained unchanged (Figure 1B), and to determine what is the environmental SXT/R391 reservoir,  
400 *i.e.* the environmental bacteria hosting these ICEs and likely being missing links in their dissemination.

401 EpicPCR, developed by Spencer *et al.* (2016) and recently improved by our group (Roman *et al.*, 2021),  
402 is a fusion PCR-based method that allows linking functional genes and phylogenetic markers (16S rRNA  
403 genes) in uncultured single cells isolated in polyacrylamide beads. Once amplified, the fusion PCR  
404 products independently obtained from millions of cells are sequenced and the operational taxonomic  
405 units (OTUs) of bacteria hosting the functional gene are identified by analyzing the 16S rDNA part of  
406 the sequences. Finally, the sorted OTUs are pointed on a 16S rDNA-based tree performed after

407 extracting and sequencing the total DNA of the microbial population also targeted by epicPCR. This  
408 method allows exploring poorly represented hosts as well as the unseen majority of non-culturable  
409 environmental bacteria.

410 For the SXT/R391 elements, epicPCR was implemented by targeting the *traB* gene, as above for qPCR.  
411 Before proceeding to massive sequencing analyses, we first validated the adequacy of our primer  
412 design for epicPCR (Table S2) by implementing this techniques at lower scale using cells recovered  
413 from the Meurthe River water upstream the WWTP, *i.e.* by cloning and sequencing the nested-PCR  
414 products obtained so as to confirm the specific fusion between SXT/R391 *traB* and 16S rDNA fragments.  
415 Of the 88 sequences obtained, all carried an expected *traB*-16S rRNA gene fusion product. Sixty-four  
416 of the corresponding 16S rDNA fragments were identified as matching the species *Nitriicola schmidtii*,  
417 a bacterium previously isolated from water of inland salt water bodies (Borsodi *et al.*, 2017). Twenty  
418 other sequences were attributed to uncultured *Aeromonas* species identical to strains isolated from  
419 aquaculture water, and the 4 remaining sequences corresponded to uncultured bacterial clones from  
420 sewage/sludge, sediment and piglet intestine. Interestingly, *Nitriicola* and *Aeromonas* species have  
421 never been identified as SXT/R391 hosts until now.

422 Once validated for fusion PCR, the primer design was further used to run epicPCR on water samples  
423 collected from the Orne and the Meurthe River waters (upstream and downstream of a WWTP  
424 discharging point), and on WWTP effluents discharged. After sequencing, data was analyzed, and  
425 phylogenetic trees were constructed as described by Roman *et al.* (2021) (Figures 3 and 4). Five  
426 confidence levels (CLs; from 1, the lower one, to 5) were defined for assigning the SXT/R391 elements  
427 to bacterial genus, depending on their presence in either technical or biological replicates (Roman *et*  
428 *al.*, 2021). SXT/R391 elements were found with different CLs in several phyla, but mainly in families  
429 belonging to the Proteobacteria phyla, as expected according to literature (Bioteau *et al.*, 2018). Despite  
430 the lack of detection of SXT/391 ICEs in the environmental *Aeromonas ssp.* isolates from our strain  
431 libraries, such elements were found in bacteria belonging to the Aeromonadaceae with the highest CLs  
432 in all tested samples covering river waters and WWTP effluents. A deeper analysis allowed identifying  
433 *Tolumonas auensis*, a non-pathogenic toluene-synthetizing bacterium previously identified in freshwater  
434 lake environments (Fischer-Romero *et al.*, 1996), as the major SXT/R391 host of this family. All the  
435 SXT/R391 hosts detected with the higher CLs belong to the  $\beta$ -,  $\gamma$ - and  $\epsilon$ -Proteobacteria phyla, such as  
436 the  $\epsilon$ -Proteobacteria Campylobacteraceae and the  $\beta$ -Proteobacteria Comamonadaceae (Figures 3 and



437 4), although all known hosts reported until now were  $\gamma$ -Proteobacteria (Bioteau *et al.*, 2018). In both  
438 ecosystems, SXT/R391 elements were found in bacteria families already mentioned in the literature,  
439 like the Enterobacteriaceae family (CLs 3 or 4 in WWTP effluents and 1 to 3 in river waters). Even if the  
440 sampling periods (September 2018 and 2019) were favorable to the environmental detection of *Vibrio*  
441 species (Crocchi *et al.*, 2001), members of this last family were not abundant enough for being detected  
442 by the 16S metagenomic approach and therefore were absent from the 16S rDNA-based phylogenetic  
443 trees (Figure 3 and 4). Considering that usually enrichment methods are necessary for recovering  
444 environmental *Vibrio* strains (Robert-Pillot *et al.*, 2002), their absence was not really surprising.  
445 Nevertheless, SXT/R391 carrying bacteria were detected with a low CL in Vibrionaceae collected in the  
446 Orne River water (Table S4). Other putative SXT/R391 hosts belonging to other taxonomic families were  
447 widely detected with a high or a low CL scores in both river ecosystems. Rather than uncertain results,  
448 and considering the relatively low abundance of SXT/R391 ICEs, the epicPCR results with a low CL  
449 score may well represent poorly abundant hosts at the edge of the detection limits of the different  
450 approaches used and should be attentively considered. Finally, an analysis of aggregated epicPCR  
451 results at the OTU level (Figure S2) demonstrated that, for the two river ecosystems studied, the WWTP  
452 effluent origins of some of the SXT/R391 hosts can be distinguished from those already present in the  
453 river upstream WWTP. This is the case, for instance, of the pathogen *Arcobacter cryaerophilus*  
454 (Campylobacteraceae) exclusively encountered in effluents and downstream river waters, as opposed  
455 to the halophilic non-pathogenic bacterium *Neptunomonas acidivorans* (Oceanospirillaceae) typically  
456 identified in rivers waters upstream and downstream the WWTP effluent discharge points. This also  
457 demonstrated that, if a quantitative impact of WWTP effluent discharges was not visible in the  
458 downstream river water regarding the abundance of SXT/R391 ICEs, a qualitative impact on the  
459 composition of SXT/R391 carrying bacterial communities could be detected. However, the results  
460 obtained also established that the environmental host-range of SXT/R391 elements differs substantially  
461 from what was observed until now, thus suggesting the existence of large environmental reservoirs that  
462 may play a pivotal role in the transmission of these ICEs from and to pathogens.

463

### 464 **3.5 Environmental significance of SXT/R391 host range broadness**

465 This work has not been focused on the cargo genes that the environmental SXT/R391 ICEs carry, *i.e.*  
466 the phenotypic traits they encode and may horizontally disseminate. Nevertheless, it could be mentioned

467 that roughly 60% and 20% of the SXT/R391 ICEs described so far carry ARGs and metal/trace element  
468 resistance genes, respectively (Bioteau *et al.*, 2018). As example, ICE*VchBan9* carries 6 ARGs (*sul2*,  
469 *floR*, *strAB*, *tetA*, *dfrA1*.) conferring resistance to 5 antibiotic classes, and 1 metal/trace element  
470 resistance gene (*zitB*). Only one fourth of these elements do not seem to encode cargo genes.  
471 Otherwise, epicPCR allows very interesting and deep metagenomic analyses but can only offer an  
472 instant picture of a microbial community. This host range does not solely reflect the sets of different  
473 bacteria meet by the SXT/R391 ICEs, but may also have a deep mechanistic implication regarding their  
474 fate and behaviors in the context of their dissemination. First, the excision of these ICEs from the  
475 chromosome of their bacterial host, the earliest step of their transfer, has been shown as being either  
476 dependent or independent of host encoded protein factors. This dependence or independence is  
477 species-specific and may impair the ability some hosts to be good SXT/R391 donors (McLeod *et al.*,  
478 2006). When present in two bacteria belonging to different species, the same SXT/R391 element can  
479 present basal transfer frequencies differing by more than one order of magnitude, and can exhibit  
480 transfer frequencies differently induce by a same concentration of an inducing molecule (Waldor *et al.*,  
481 1996; Beaber *et al.*, 2004). Beyond these phenomena, SXT/R391 ICEs can carry a huge allelic variety  
482 of core genes and promoters (Bioteau *et al.*, 2018), that again multiplies the kinds of behavior they can  
483 have in a same species. The study of another ICE, Tn*916*, has shown that the same element can be  
484 very stable in a species, while highly unstable in another one. This instability has been linked to a very  
485 high excision frequency of the ICE and its subsequent lost during cell divisions, likely because of the  
486 misregulation of the promoters controlling the ICE recombination module (Scott *et al.*, 1988). Indeed, as  
487 the consensus sequences of bacterial promoters vary between species, a same gene can be differently  
488 expressed in two species because of the distinct recognition efficiencies of its promoter by the sigma  
489 factors of the bacteria where it has transferred (Haugen *et al.*, 2008). Finally, as others MGEs, SXT/R391  
490 ICEs carry addiction modules, actually toxin-antitoxin systems, that promote the maintenance of the  
491 acquired elements (Wozniak and Waldor, 2009). Any mutation or misregulation in the expression of  
492 such an addiction module can lead in the ICE loss at a high frequency, or may impact on the viability of  
493 the bacterial cells carrying an MGE. All in all, these different phenomena draw a complex picture where  
494 the same SXT/R391 element can excise, transfer, integrate and maintained over cell divisions at  
495 different frequencies when it is harbored by hosts belonging to different bacterial species. Thus, if  
496 SXT/R391 elements are present in a variety of hosts within a microbial community, they likely behave

497 differently, may have many fates, and may variously respond to the same stimulating molecules.  
498 Therefore, there may be as much behaviors of SXT/R391 ICEs than combinations of such an ICE within  
499 a particular host. The host range and abundance of such MGEs in a microbial community should then  
500 have to be thought as ever-changing equilibriums rather than fixed characteristics. This again  
501 complexifies the role these ICEs could have in disseminating the cargo genes they encode, and augurs  
502 new and exciting studies to better define the role plays by these elements in ecosystems.

503

504

#### 505 **4 Conclusion**

506 The design of new sets of primers/probe targeting the conserved *traB* gene of SXT/R391 elements was  
507 a key step towards the quantification and the taxonomic assignment of these ICEs in aquatic  
508 environments. All in all, SXT/R391 elements appeared to be likely ubiquitous and much more abundant  
509 than anticipated (from  $10^{-6}$  to  $10^{-3}$  ICE per 16S rRNA gene), especially in slightly salty or brackish water  
510 that may well represent a key environmental parameter responsible for the persistence of SXT/R391  
511 hosts, and thus a putative environmental hot spot for SXT/R391-based ARG propagation. Thanks to the  
512 epicPCR, the environmental host range of SXT/R391 hosts could be explored without *a priori* beyond  
513 the limits of the classically described *Vibrio spp.* SXT/R391 elements were rather found in bacterial  
514 families (Aeromonadaceae, Campylobacteraceae Comamonadaceae) and species (*T. auensis*; *A.*  
515 *cryaerophilus*) that were until now not mentioned in literature. Considering the fact that *Vibrio spp.* were  
516 classically reported and monitored as the most frequent hosts of SXT/R391 ICEs, this work shows that,  
517 actually, great care should be given to the unseen majority represented by the environmental bacteria  
518 hosting SXT/R391 elements that likely are the main sources of these MGEs and their carried ARGs  
519 towards pathogens. In other words, following the One Health concept, the propagation of MGEs/ARGs  
520 should also be investigated in non-clinical isolates when it comes to understanding the connection  
521 between environmental reservoirs, pollution sources (e.g. WWTP effluents) and human/animal  
522 exposure.

523

524

#### 525 **5 Acknowledgements**

526 The authors want to thank B. Gassilloud, A. Wilhelm (MALDI-TOF platform, Nancy Hydrology Laboratory  
527 (LHN), Anses), T. Morin, M. Danion (Ploufragan-Plouzané-Niort Laboratory, Anses), M. Montargès-  
528 Peleltier, C. Gaultier (LIEC, CNRS, Université de Lorraine), T. Lecocq, Y. Ledore (URAFPA, Université  
529 de Lorraine), C. Demangeclaude and S. Trévarin (Véolia, Maxéville), the Métropole du Grand Nancy  
530 (intercommunal structure centered on Nancy City) for their help in sampling and fish associated strains  
531 isolation and characterization.

532

533

#### 534 **6 Funding sources**

535 VLR has been supported by a Short-Term Scientific Mission Grant from the Nereus COST Action  
536 ES1403. CM and XB wish to thank the French Agency for Food, Environmental and Occupational Health  
537 & Safety (Project “AquaResist,” PNREST Anses, 2018/1/052), the Carnot Institute ICÉEL, and the Zone  
538 Atelier Moselle for research funding. SB was supported by grants from the European Maritime and  
539 Fisheries Fund (project Resist3A), Ecoantibio 1 plan (French Ministry of Agriculture; project Antibiofish),  
540 and the Service des études médicales of Électricité de France (EDF). MV is funded by Academy of  
541 Finland.

542

#### 543 **7 CRediT author statement**

544 **Veronica L. Roman:** Investigation, Methodology, Visualization, Writing - original draft. **Christophe**  
545 **Merlin:** Conceptualization, Data curation, Funding acquisition, Methodology, Project administration,  
546 Supervision, Writing - original draft, Writing - review & editing. **Sandrine Baron:** Data curation, Funding  
547 acquisition, Writing - review & editing. **Emeline Larvor:** Investigation. **Laetitia Le Devendec:**  
548 Investigation. **Marko P. J. Virta:** Data curation, Methodology, Writing - review & editing. **Xavier**  
549 **Bellanger:** Conceptualization, Data curation, Funding acquisition, Methodology, Supervision, Writing -  
550 original draft; Writing - review & editing.

551

552

#### 553 **8 Competing Interests**

554 None to declare

555

556

557 **8 References**

558 Aarestrup, F. M., Seyfarth, A. M., Emborg, H. D., Pedersen, K., Hendriksen, R. S., Bager, F., 2001.

559 Effect of abolishment of the use of antimicrobial agents for growth promotion on occurrence of

560 antimicrobial resistance in fecal enterococci from food animals in Denmark. *Antimicrobial agents and*

561 *chemotherapy*, 45(7), 2054-2059. <https://doi.org/10.1128/AAC.45.7.2054-2059.2001>

562 Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., Lipman, D. J., 1997.

563 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic*

564 *acids research*, 25(17), 3389-3402. <https://doi.org/10.1093/nar/25.17.3389>

565 Andersson, D. I., Hughes, D., 2014. Microbiological effects of sublethal levels of antibiotics. *Nature*

566 *Reviews Microbiology*, 12(7), 465-478. <https://doi.org/10.1038/nrmicro3270>

567 Andrews S. FastQC. Babraham Bioinformatics [cited on July 6, 2020]. Accessible on:

568 <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

569 Aubertreau, E., Stalder, T., Mondamert, L., Ploy, M. C., Dagot, C., Labanowski, J., 2017. Impact of

570 wastewater treatment plant discharge on the contamination of river biofilms by pharmaceuticals and

571 antibiotic resistance. *Science of the Total Environment*, 579, 1387-1398.

572 <https://doi.org/10.1016/j.scitotenv.2016.11.136>

573 Badhai, J., Das, S. K., 2016. Characterization of three novel SXT/R391 integrating conjugative

574 elements ICEMfulnd1a and ICEMfulnd1b, and ICEMprChn1 identified in the genomes of

575 *Marinomonas fungiae* JCM 18476T and *Marinomonas profundimaris* strain D104. *Frontiers in*

576 *microbiology*, 7, 1896. <https://doi.org/10.3389/fmicb.2016.01896>

577 Balado, M., Lemos, M. L., Osorio C, R., 2013. Integrating conjugative elements of the SXT/R391

578 family from fish-isolated *Vibrios* encode restriction–modification systems that confer resistance to

579 bacteriophages. *FEMS microbiology ecology*, 83(2), 457-467. <https://doi.org/10.1111/1574->

580 6941.12007

581 Baron, S., Lesne, J., Moore, S., Rossignol, E., Rebaudet, S., Gazin, P., ... Piarroux, R., 2013. No  
582 evidence of significant levels of toxigenic *V. cholerae* O1 in the Haitian aquatic environment during the  
583 2012 rainy season. PLoS currents, 5.  
584 <https://dx.doi.org/10.1371%2Fcurrents.outbreaks.7735b392bdcb749baf5812d2096d331e>

585 Baron, S., Granier, S. A., Larvor, E., Jouy, E., Cineux, M., Wilhelm, A., ... Chauvin, C., 2017.  
586 *Aeromonas* diversity and antimicrobial susceptibility in freshwater—an attempt to set generic  
587 epidemiological cut-off values. Frontiers in microbiology, 8, 503.  
588 <https://doi.org/10.3389/fmicb.2017.00503>

589 Beaber, J. W., Burrus, V., Hochhut, B., & Waldor, M. K., 2002a. Comparison of SXT and R391, two  
590 conjugative integrating elements: definition of a genetic backbone for the mobilization of resistance  
591 determinants. Cellular and Molecular Life Sciences CMLS, 59(12), 2065-2070.  
592 <https://doi.org/10.1007/s000180200006>

593 Beaber, J. W., Hochhut, B., & Waldor, M. K., 2002b. Genomic and functional analyses of SXT, an  
594 integrating antibiotic resistance gene transfer element derived from *Vibrio cholerae*. Journal of  
595 Bacteriology, 184(15), 4259-4269. <https://doi.org/10.1128/JB.184.15.4259-4269.2002>

596 Beaber, J. W., Hochhut, B., & Waldor, M. K., 2004. SOS response promotes horizontal dissemination  
597 of antibiotic resistance genes. Nature, 427(6969), 72-74. <https://doi.org/10.1038/nature02241>

598 Benami, M., Gross, A., Herzberg, M., Orlofsky, E., Vonshak, A., Gillor, O., 2013. Assessment of  
599 pathogenic bacteria in treated graywater and irrigated soils. Science of the total environment, 458,  
600 298-302. <https://doi.org/10.1016/j.scitotenv.2013.04.023>

601 Bioteau, A., Durand, R., Burrus, V., 2018. Redefinition and unification of the SXT/R391 family of  
602 integrative and conjugative elements. Applied and environmental microbiology, 84(13).  
603 <https://doi.org/10.1128/AEM.00485-18>

604 Borsodi, A. K., Korponai, K., Schumann, P., Spröer, C., Felföldi, T., Márialigeti, K., ... Tóth, E., 2017.  
605 *Nitriicola alkalilacustris* sp. nov. and *Nitriicola schmidtii* sp. nov., alkaliphilic bacteria isolated from  
606 soda pans, and emended description of the genus *Nitriicola*. International Journal of Systematic and

607 Evolutionary Microbiology, 67(12), 5159-5164. <https://doi.org/10.1099/ijsem.0.002437>

608 Calderón-Franco, D., van Loosdrecht, M. C., Abeel, T., & Weissbrodt, D. G., 2020. Free-floating  
609 extracellular DNA: systematic profiling of mobile genetic elements and antibiotic resistance from  
610 wastewater. *Water Research*, 116592.

611 Coetzee, J. N., Datta, N., Hedges, R. W., 1972. R factors from *Proteus rettgeri*. *Microbiology*, 72(3),  
612 543-552. <https://doi.org/10.1099/00221287-72-3-543>

613 Collignon, P., Beggs, J. J., Walsh, T. R., Gandra, S., Laxminarayan, R., 2018. Anthropological and  
614 socioeconomic factors contributing to global antimicrobial resistance: a univariate and multivariable  
615 analysis. *The Lancet Planetary Health*, 2(9), e398-e405. <https://doi.org/10.1016/S2542->  
616 5196(18)30186-4

617 Courvalin, P., 2005. Antimicrobial drug resistance:" Prediction is very difficult, especially about the  
618 future". *Emerging Infectious Diseases*, 11(10), 1503. <https://dx.doi.org/10.3201%2F1110.051014>

619 Croci, L., Serratore, P., Cozzi, L., Stacchini, A., Milandri, S., Suffredini, E., Toti, L., 2001. Detection of  
620 *Vibrionaceae* in mussels and in their seawater growing area. *Letters in Applied Microbiology*, 32(1),  
621 57-61. <https://doi.org/10.1111/j.1472-765X.2001.00855.x>

622 Davies, J., 2007. Microbes have the last word: A drastic re-evaluation of antimicrobial treatment is  
623 needed to overcome the threat of antibiotic-resistant bacteria. *EMBO reports*, 8(7), 616-621.  
624 <https://doi.org/10.1038/sj.embor.7401022>

625 Davies, J., Davies, D., 2010. Origins and evolution of antibiotic resistance. *Microbiology and molecular*  
626 *biology reviews*, 74(3), 417-433. <https://doi.org/10.1128/MMBR.00016-10>

627 De la Cruz Barrón, M., Merlin, C., Guilloteau, H., Montargès-Pelletier, E., Bellanger, X., 2018.  
628 Suspended materials in river waters differentially enrich class 1 integron-and IncP-1 plasmid-carrying  
629 bacteria in sediments. *Frontiers in microbiology*, 9, 1443. <https://doi.org/10.3389/fmicb.2018.01443>

630 Edgar, R. C., 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*,  
631 26(19), 2460-2461. <https://doi.org/10.1093/bioinformatics/btq461>

632 Fischer-Romero, C., Tindall, B. J., Jüttner, F., 1996. *Tolomonas auensis* gen. nov., sp. nov., a toluene-  
633 producing bacterium from anoxic sediments of a freshwater lake. International Journal of Systematic  
634 and Evolutionary Microbiology, 46(1), 183-188. <https://doi.org/10.1099/00207713-46-1-183>

635 Furuya, E. Y., Lowy, F. D., 2006. Antimicrobial-resistant bacteria in the community setting. Nature  
636 Reviews Microbiology, 4(1), 36-45. <https://doi.org/10.1038/nrmicro1325>

637 Haugen, S. P., Ross, W., Gourse, R. L., 2008. Advances in bacterial promoter recognition and its  
638 control by factors that do not bind DNA. Nature Reviews Microbiology, 6(7), 507-519.  
639 <https://doi.org/10.1038/nrmicro1912>

640 Hochhut, B., Beaber, J. W., Woodgate, R., Waldor, M. K., 2001. Formation of chromosomal tandem  
641 arrays of the SXT element and R391, two conjugative chromosomally integrating elements that share  
642 an attachment site. Journal of bacteriology, 183(4), 1124-1132. [https://doi.org/10.1128/JB.183.4.1124-](https://doi.org/10.1128/JB.183.4.1124-1132.2001)  
643 [1132.2001](https://doi.org/10.1128/JB.183.4.1124-1132.2001)

644 Hochhut, B., Waldor, M. K., 1999. Site-specific integration of the conjugal *Vibrio cholerae* SXT element  
645 into *prfC*. Molecular microbiology, 32(1), 99-110. <https://doi.org/10.1046/j.1365-2958.1999.01330.x>

646 Hultman, J., Tamminen, M., Pärnänen, K., Cairns, J., Karkman, A., Virta, M., 2018. Host range of  
647 antibiotic resistance genes in wastewater treatment plant influent and effluent. FEMS microbiology  
648 ecology, 94(4), fiy038. <https://doi.org/10.1093/femsec/fiy038>

649 Knapp, C. W., Dolfing, J., Ehlert, P. A., Graham, D. W., 2010. Evidence of increasing antibiotic  
650 resistance gene abundances in archived soils since 1940. Environmental science & technology, 44(2),  
651 580-587. <https://doi.org/10.1021/es901221x>

652 Lei, C. W., Zhang, A. Y., Wang, H. N., Liu, B. H., Yang, L. Q., Yang, Y. Q., 2016. Characterization of  
653 SXT/R391 integrative and conjugative elements in *Proteus mirabilis* isolates from food-producing  
654 animals in China. Antimicrobial agents and chemotherapy, 60(3), 1935-1938.  
655 <https://doi.org/10.1128/AAC.02852-15>

656 Letunic, I., Bork, P., 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments.  
657 Nucleic acids research, 47(W1), W256-W259. <https://doi.org/10.1093/nar/gkz239>



658 Li, Y., Li, Y., Fernandez Crespo, R., Leanse, L. G., Langford, P. R., & Bossé, J. T., 2018.  
659 Characterization of the *Actinobacillus pleuropneumoniae* SXT-related integrative and conjugative  
660 element ICE*ApI2* and analysis of the encoded FloR protein: Hydrophobic residues in transmembrane  
661 domains contribute dynamically to florfenicol and chloramphenicol efflux. *Journal of Antimicrobial*  
662 *Chemotherapy*, 73(1), 57-65. <https://doi.org/10.1093/jac/dkx342>

663 McLeod, S. M., Burrus, V., Waldor, M. K., 2006. Requirement for *Vibrio cholerae* integration host  
664 factor in conjugative DNA transfer. *Journal of bacteriology*, 188(16), 5704-5711.  
665 <https://doi.org/10.1128/JB.00564-06>

666 Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., ... Lopez, R., 2019. The  
667 EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic acids research*, 47(W1), W636-  
668 W641. <https://doi.org/10.1093/nar/gkz268>

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

671 Martinez, J. L., 2009. Environmental pollution by antibiotics and by antibiotic resistance determinants.  
672 *Environmental pollution*, 157(11), 2893-2902. <https://doi.org/10.1016/j.envpol.2009.05.051>

673 Pembroke, J. T., Piterina, A. V., 2006. A novel ICE in the genome of *Shewanella putrefaciens* W3-18-  
674 1: comparison with the SXT/R391 ICE-like elements. *FEMS microbiology letters*, 264(1), 80-88.  
675 <https://doi.org/10.1111/j.1574-6968.2006.00452.x>

676 Price, M. N., Dehal, P. S., & Arkin, A. P., 2009. FastTree: computing large minimum evolution trees  
677 with profiles instead of a distance matrix. *Molecular biology and evolution*, 26(7), 1641-1650.  
678 <https://doi.org/10.1093/molbev/msp077>

679 Robert-Pillot, A., Baron, S., Lesne, J., Fournier, J. M., Quilici, M. L., 2002. Improved specific detection  
680 of *Vibrio cholerae* in environmental water samples by culture on selective medium and colony  
681 hybridization assay with an oligonucleotide probe. *FEMS Microbiology Ecology*, 40(1), 39-46.  
682 <https://doi.org/10.1111/j.1574-6941.2002.tb00934.x>

683 Rocha, J., Cacace, D., Kampouris, I., Guilloteau, H., Jäger, T., Marano, R. B., ... Cytryn, E., 2020.

684 Inter-laboratory calibration of quantitative analyses of antibiotic resistance genes. Journal of  
685 Environmental Chemical Engineering, 8(1), 102214. <https://doi.org/10.1016/j.jece.2018.02.022>

686 Roman VL, Merlin C, Virta MPJ, Bellanger X. EpicPCR 2.0: Technical and methodological  
687 improvement of a recent single-cell genomic approach. (under submission)

688 Sato, J. L., Fonseca, M. R., Cerdeira, L. T., Tognim, M. C., Sincero, T. C., do Amaral, M. C. N., ... &  
689 Galhardo, R. S., 2020. Genomic Analysis of SXT/R391 Integrative Conjugative Elements From  
690 *Proteus mirabilis* Isolated in Brazil. *Frontiers in Microbiology*, 11.  
691 <https://dx.doi.org/10.3389/fmicb.2020.571472>

692 Scott, J. R., Kirchman, P. A., Caparon, M. G., 1988. An intermediate in transposition of the conjugative  
693 transposon Tn916. *Proceedings of the National Academy of Sciences*, 85(13), 4809-4813.  
694 <https://doi.org/10.1073/pnas.85.13.4809>

695 Shinde, A. H., Raval, I. H., & Haldar, S., 2020. SXT *int* harboring bacteria as effective indicators to  
696 determine high-risk reservoirs of multiple antibiotic resistance in different aquatic environments of  
697 western coast of Gujarat, India. *Ecological Indicators*, 113, 106143.  
698 <https://doi.org/10.1016/j.ecolind.2020.106143>

699 Spagnoletti, M., Ceccarelli, D., Rieux, A., Fondi, M., Taviani, E., Fani, R., ... Balloux, F., 2014.  
700 Acquisition and evolution of SXT-R391 integrative conjugative elements in the seventh-pandemic  
701 *Vibrio cholerae* lineage. *MBio*, 5(4). <https://doi.org/10.1128/mBio.01356-14>

702 Spencer, S. J., Tamminen, M. V., Preheim, S. P., Guo, M. T., Briggs, A. W., Brito, I. L., ... & Alm, E. J.,  
703 2016. Massively parallel sequencing of single cells by epicPCR links functional genes with  
704 phylogenetic markers. *The ISME journal*, 10(2), 427-436. <https://doi.org/10.1038/ismej.2015.124>

705 Waldor, M. K., Tschäpe, H. E. L. M. U. T., & Mekalanos, J. J., 1996. A new type of conjugative  
706 transposon encodes resistance to sulfamethoxazole, trimethoprim, and streptomycin in *Vibrio cholerae*  
707 O139. *Journal of bacteriology*, 178(14), 4157-4165. <https://doi.org/10.1128/jb.178.14.4157-4165.1996>

708 World Health Organization. Cholera. 2019 [cited on July 6, 2020]. Accessible on:  
709 <https://www.who.int/fr/news-room/fact-sheets/detail/cholera>

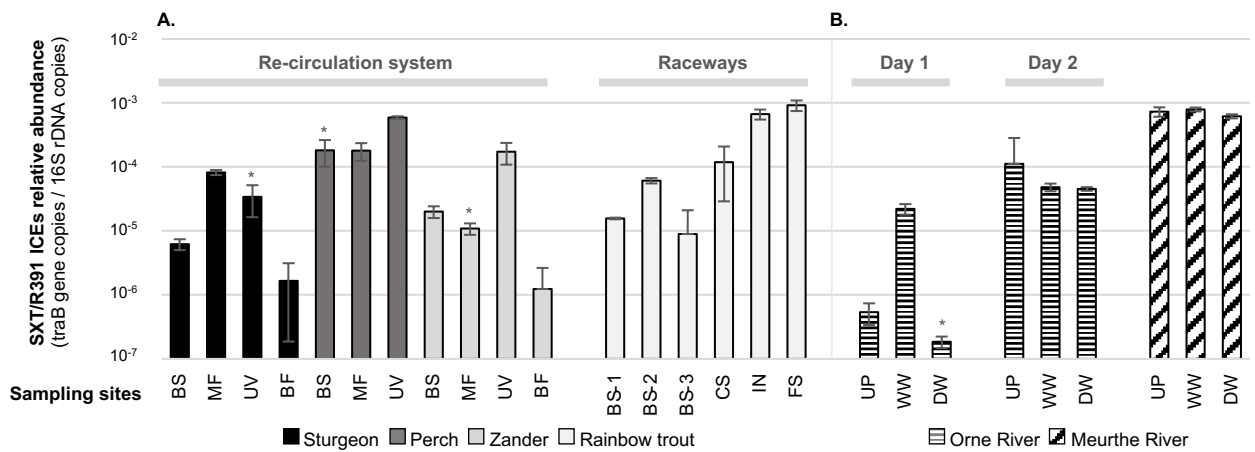
710 World Health Organization. Global Health Observatory (GHO) - Number of deaths due to cholera 2020  
711 [cited on July 6, 2020]. Accessible on:  
712 [https://www.who.int/gho/epidemic\\_diseases/cholera/situation\\_trends\\_deaths/en/](https://www.who.int/gho/epidemic_diseases/cholera/situation_trends_deaths/en/)

713 Wozniak, R. A., Fouts, D. E., Spagnoletti, M., Colombo, M. M., Ceccarelli, D., Garriss, G., ... & Waldor,  
714 M. K., 2009. Comparative ICE genomics: insights into the evolution of the SXT/R391 family of ICEs.  
715 *PLoS Genet*, 5(12), e1000786.

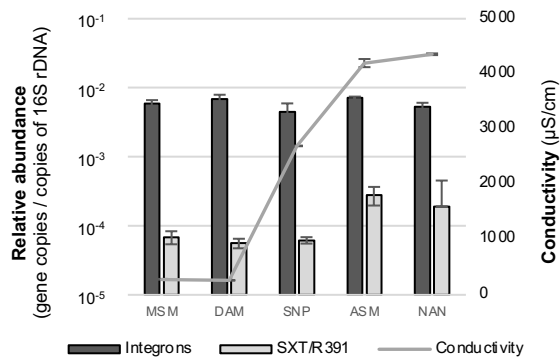
716 Wozniak, R. A., Waldor, M. K., 2009. A Toxin–Antitoxin System Promotes the Maintenance of an  
717 Integrative Conjugative Element. *PLoS Genet* 5(3): e1000439. doi:10.1371/journal.pgen.1000439

718 Xiong, W., Wang, Y., Sun, Y., Ma, L., Zeng, Q., Jiang, X., ... Zhang, T., 2018. Antibiotic-mediated  
719 changes in the fecal microbiome of broiler chickens define the incidence of antibiotic resistance genes.  
720 *Microbiome*, 6(1), 1-11. <https://doi.org/10.1186/s40168-018-0419-2>

721 Zhang, J., Kobert, K., Flouri, T., Stamatakis, A., 2014. PEAR: a fast and accurate Illumina Paired-End  
722 reAd mergeR. *Bioinformatics*, 30(5), 614-620. <https://doi.org/10.1093/bioinformatics/btt593>

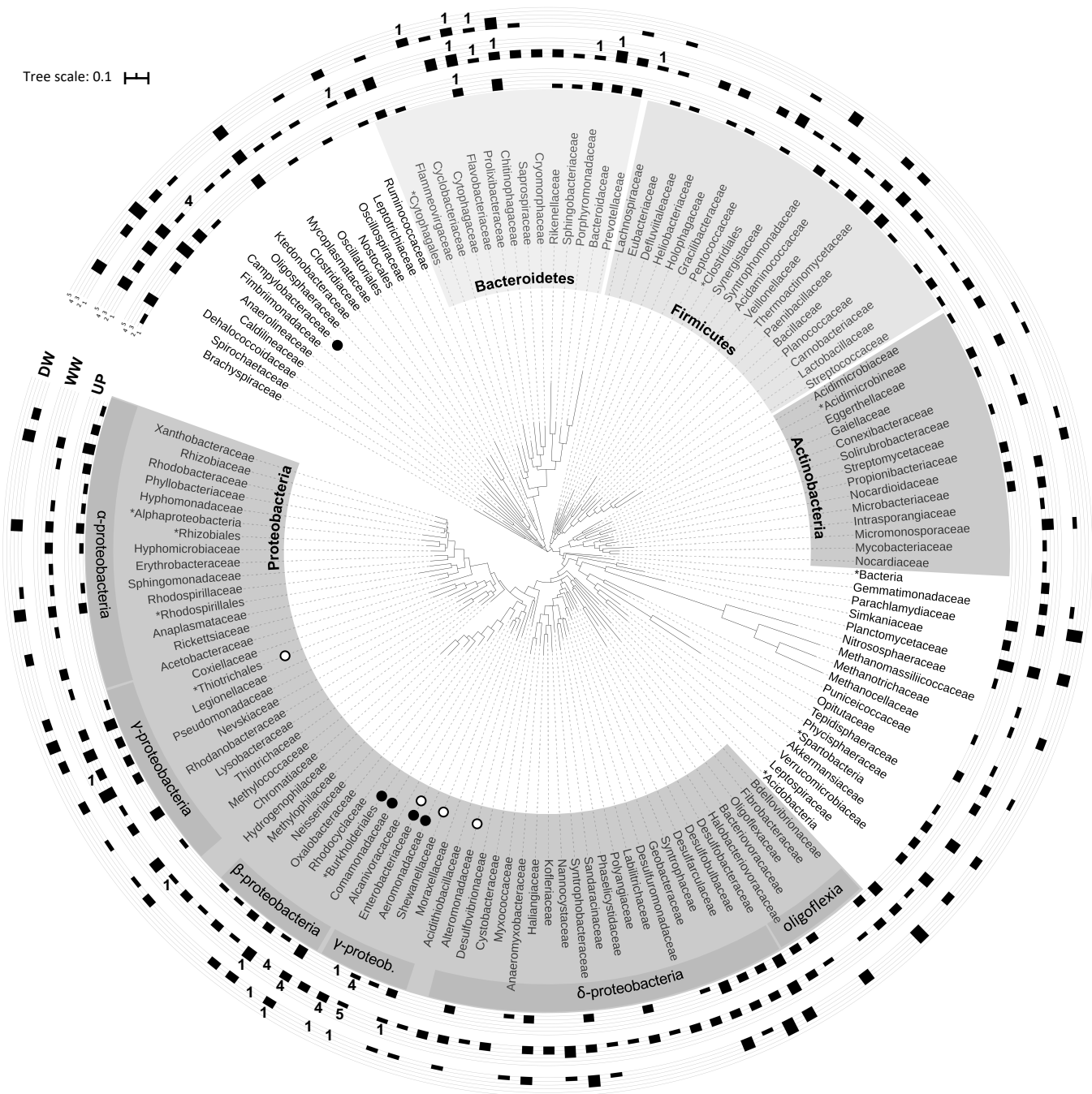


**Figure 1.** Relative abundance of SXT/R391 ICEs in (A) water, biofilm and fish samples from two fish farming facilities and (B) water samples from the Orne and the Meurthe Rivers. BS, water from the basin; MF, water from the mechanical filtering system; UV, water collected after the UV treatment system; BF, inner tank surface biofilm; SC, fish stomach contents; IN, fish intestine content; FS, fish feces; UP, river water collected upstream a WWTP; WW, WWTP effluents; DW, river water collected downstream a WWTP. Error bars indicate the standard error of the mean (n=3 except for BS, MF, UV and BF where n=1 and for which error bars indicate the standard error of the qPCR triplicates). Asterisks indicate detected values close to the qPCR limit of quantification



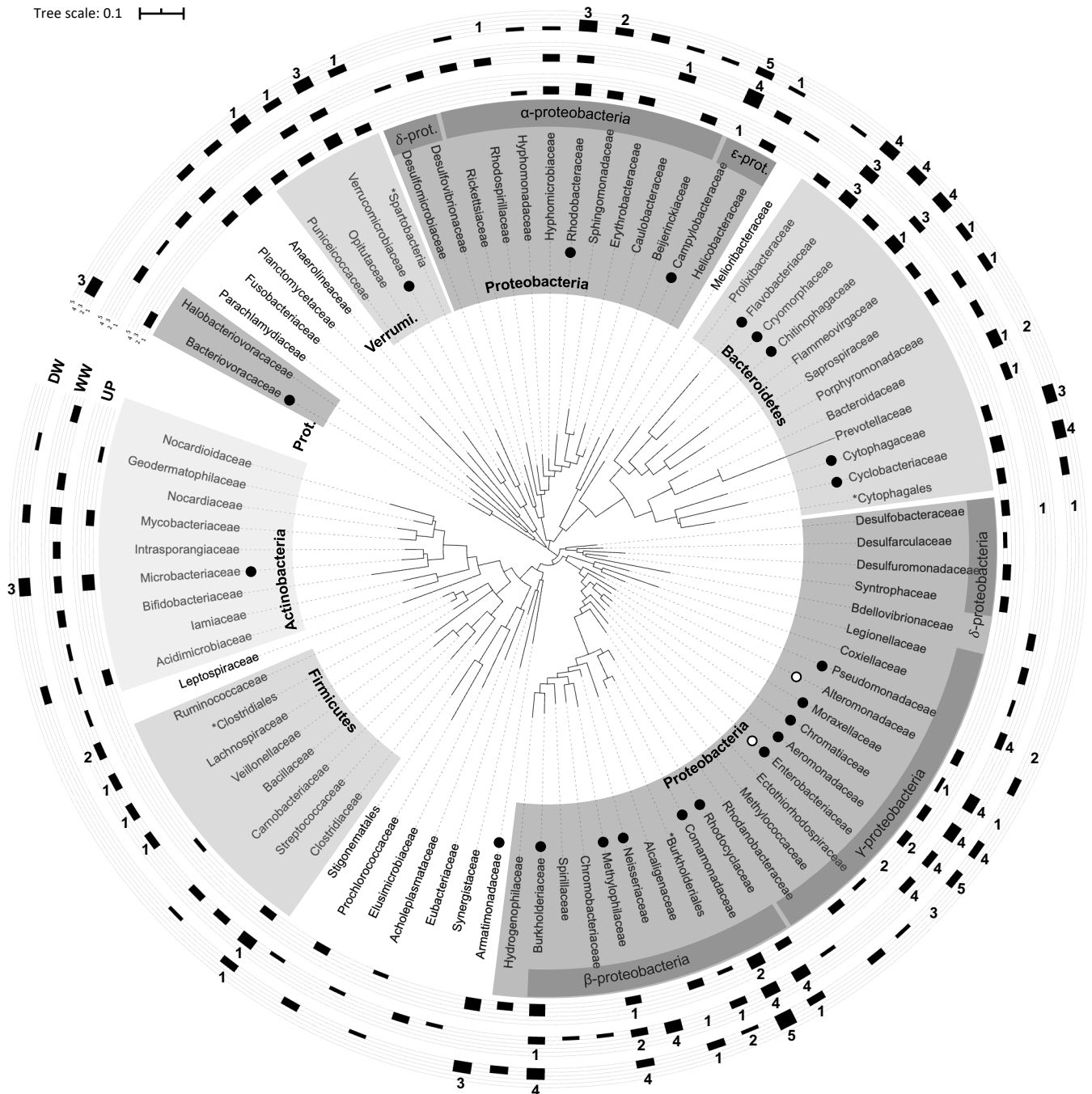
**Figure 2.** Relative abundance of SXT/R391 ICEs and class 1 integrons in water samples from the Meurthe River and conductivity measurements. Samples were collected in various cities along the Meurthe River continuum, from upstream to downstream: MSM, Mont-sur-Meurthe; DAM, Damelevières; SNP, Saint-Nicolas-de-Port; ASM, Art-sur-Meurthe; NAN, Nancy. All the sampling sites are located upstream of the WWTP of the Nancy City area and its discharging point into the Meurthe River. Error bars indicate the standard error of the mean (n=3).

Tree scale: 0.1



**Figure 3.** Host range of SXT/R391 ICEs in the Orne River ecosystem identified by epicPCR. WW: wastewater treatment plant effluents; UP: river water collected upstream of the WWTP; DW: river water collected downstream of the WWTP. The phylogenetic tree was constructed from family-level OTUs based on 16S rRNA gene amplicon sequences. The five confidence levels (from 1, the lowest, to 5) of the presence of SXT/R391 elements in a bacteria family are indicated by numbers. Open and full black circles indicate taxonomic families where SXT/R391 elements have been previously detected or detected in the present work with a high confidence level (CL $\geq$ 3), respectively. The confidence levels are not related to the black rectangles in the UP, WW and DW rings that display the relative abundances of 16S rDNA reads for the considered family. The numbers at the edge of the rings indicate the Log10 values of the calculation (sum of OTU reads from a family / total reads of a sample) \* 100 000 expressing the relative OTU read abundances. Asterisks are for unclassified families.

Tree scale: 0.1



**Figure 4.** Host range of SXT/R391 ICEs in the Meurthe River ecosystem identified by epicPCR. WW: wastewater treatment plant effluents; UP: river water collected upstream of the WWTP; DW: river water collected downstream of the WWTP. The phylogenetic tree was constructed from family-level OTUs based on 16S rRNA gene amplicon sequences. The five confidence levels (from 1, the lowest, to 5) of the presence of SXT/R391 elements in a bacteria family are indicated by numbers. Open and full black circles indicate taxonomic families where SXT/R391 elements have been previously detected or detected in the present work with a high confidence level (CL $\geq$ 3), respectively. The confidence levels are not related to the black rectangles in the UP, WW and DW rings that display the relative abundances of 16S rDNA reads for the considered family. The numbers at the edge of the rings indicate the Log<sub>10</sub> values of the calculation (sum of OTU reads from a family / total reads of a sample) \* 100 000 expressing the relative OTU read abundances. Asterisks are for unclassified families.

1 **SUPPLEMENTARY MATERIAL**

2

3 **Abundance and environmental host range of the SXT/R391 ICEs in aquatic environmental**  
4 **communities.**

5

6 Veronica L. Roman<sup>1</sup>, Christophe Merlin<sup>1</sup>, Sandrine Baron<sup>2</sup>, Emeline Larvor<sup>2</sup>, Laetitia Ledevendec<sup>2</sup>,

7 Marko P.J. Virta<sup>3</sup>, Xavier Bellanger<sup>1\*</sup>

8

9 <sup>1</sup>LCPME, CNRS, Université de Lorraine, Nancy, France

10 <sup>2</sup>Mycoplasmology-Bacteriology Unit, Ploufragan-Plouzané Laboratory, French Agency for Food,

11 Environmental and Occupational Health and Safety (Anses), Ploufragan, France

12 <sup>3</sup>Department of Microbiology, University of Helsinki, Helsinki, Finland

13

14

15 \* corresponding author: [xavier.bellanger@univ-lorraine.fr](mailto:xavier.bellanger@univ-lorraine.fr)

16

17



18 **Table S2:** Primers used in this study for qPCR (a), epicPCR (b), 16S rRNA gene amplification for  
 19 NGS (c) and plasmid sequencing (d).

20

Name	a	b	c	d	Sequence (5' - 3')	Targeted gene	Reference
traB <sub>SXT-1</sub>	X			X	CTGCATTRCTGAGGATTCAA	<i>traB</i> (SXT/R391 elements)	This work
traB <sub>SXT-2</sub>	X	X		X	CGGCGATCAAGAAACACTCT	<i>traB</i> (SXT/R391 elements)	This work
traB <sub>SXT-p</sub>	X				TACCYAACCGATTYAGAGCK	<i>traB</i> probe (SXT/R391 elements)	This work
int <sub>SXT-1</sub>	X				TGAAAAACAATTTTGCCCTTG	<i>int</i> (SXT/R391 elements)	This work
int <sub>SXT-2</sub>	X				ATCAGTTAGCTGGCTCGATG	<i>int</i> (SXT/R391 elements)	This work
int <sub>SXT-p</sub>	X				CWCGRTCRCGCTTTTACTACG	<i>int</i> probe (SXT/R391 elements)	This work
intI1LC5	X				GATCGGTCTGAATGCGTGT	<i>intI1</i> (class 1 integrons)	Götz <i>et al.</i> , (1996)
intI1LC1	X				GCCTTGATGTTACCGAGAG	<i>intI1</i> (class 1 integrons)	Götz <i>et al.</i> , (1996)
338F	X				CCTACGGGAGGCAGCAG	16S rRNA	Muyzer <i>et al.</i> , (1993)
518R	X				ATTACCGCGGCTGCTGG	16S rRNA	Muyzer <i>et al.</i> , (1993)
traB <sub>SXT-1_519R</sub>		X			GWATTACCGCGGCKGCTGCTGCATTRCTG AGGATTCAA	<i>traB</i> epicPCR linker	This work
1492R		X			GGTTACCTTGTACGACTT	16S rRNA	Lane (1991)
traB <sub>SXT-3_comp-inv</sub>		X			MGCTCTRAATCGGTRGGTA	<i>traB</i> epicPCR nested	This work
806R			X		GGACTACHVGGGTWTCTAAT	16S rRNA	Tremblay <i>et al.</i> , (2015)
515F			X		GTGCCAGCMGCCGCGGTAA	16S rRNA	Tremblay <i>et al.</i> , (2015)
Illum_785R_1		X			GTGACTGGAGTTCAGACGTGTGCTCTTCC GATCTGACTACHVGGGTATCTAATCC	16S rRNA reverse primer with Illumina TruSeq adapter	Hultman <i>et al.</i> , (2018)
Illum_785R_2		X			GTGACTGGAGTTCAGACGTGTGCTCTTCC GATCTaGACTACHVGGGTATCTAATCC	16S rRNA reverse primer with Illumina TruSeq adapter	Hultman <i>et al.</i> , (2018)
Illum_785R_3		X			GTGACTGGAGTTCAGACGTGTGCTCTTCC GATCTtGACTACHVGGGTATCTAATCC	16S rRNA reverse primer with Illumina TruSeq adapter	Hultman <i>et al.</i> , (2018)
Illum_785R_4		X			GTGACTGGAGTTCAGACGTGTGCTCTTCC GATCTcgagtGACTACHVGGGTATCTAATC C	16S rRNA reverse primer with Illumina TruSeq adapter	Hultman <i>et al.</i> , (2018)
SXT-3_comp-inv-Illumina		X			ATCTACACTCTTCCCTACACGACGCTCTT CCGATCTMGCTCTRAATCGGTRGGTA	<i>traB</i> epicPCR nested primer with Illumina TruSeq adapter	This work
U519R-block10		X			TTTTTTTTTGWATTACCGCGGCKGCTG/3S pC3/	epicPCR unfused products	Spencer <i>et al.</i> , (2016)
U519F-block10		X			TTTTTTTTTTCAGCMGCCGCGGTAATWC/3 SpC3/	epicPCR unfused products	Spencer <i>et al.</i> , (2016)
SP6				X	GATTTAGGTGACACTATAG	Universal sequencing primer	

21

22

23

24

25

26 **Table S3:** Nucleotide sequence polymorphism of the 34 cloned *traB* gene fragments amplified using  
 27 total DNA extracted from Meurthe River water.

Seq. type <sup>a</sup>	Position in base pairs (from 1 to 81)						Sequence occurrence	Identical reference sequence (accession number)
	8th	30th	47th	51st	53th	56th		
1	G	A	T	G	G	C	17	Not yet listed in databases
2	G	A	T	G	G	T	11	ICE <i>Vch</i> CHN4210 (KT151662.1)
3	A	A	T	G	G	C	9	ICE SXT <sub>MO10</sub> (AY055428.1)
4	G	A	C	G	G	T	6	Not yet listed in databases
5	G	G	T	A	T	T	3	Not yet listed in databases

28

29 These 81 bp-fragments were amplified using primers *traB*-SXT1 and *traB*-SXT2 listed in Table S2.

30 <sup>a</sup>: the reference sequence from ICE SXT<sub>MO10</sub> (AY055428.1) (here, sequence type 3) is:  
 31 ctgcattactgaggattcaaaaggaagcc**at**tttacc**ca**accgatttagag**cg**gacatcaaagagtg**tt**ctt**g**atcgccg. The varying  
 32 nucleotides are presented as bold characters.

33

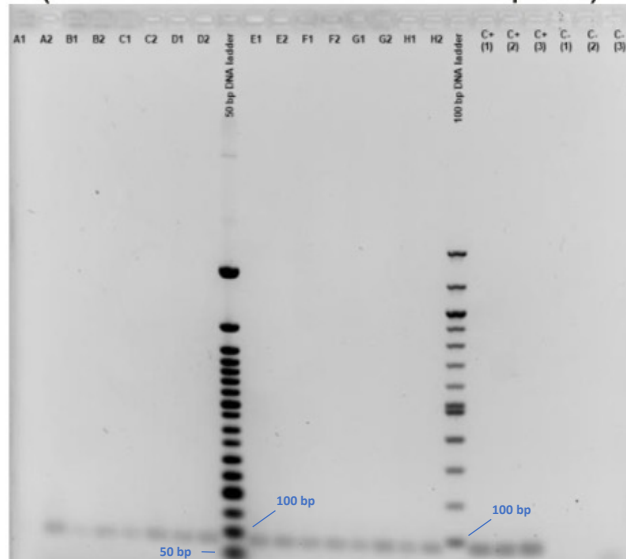
34 ctgcatt**g**ctgaggattcaaaaggaagcc**g**tttacc**ca**accgatttaga**act**gatatcaaagagtg**tt**ctt**g**atcgccg

35 **Table S4:** EpicPCR results and their associated level of confidence for the Meurthe and the Orne  
 36 Rivers when reads from 16S sRNA gene sequencing were lower than 10 per OTU, and therefore not  
 37 represented in phylogenetic trees.

River	Sampling site	Level of confidence	Family
Meurthe	UP	1	Campylobacteraceae
		2	Enterobacteriaceae ○
		1	Idiomarinaceae ○
		2	Oceanospirillaceae ●
	WW	1	Alcaligenaceae
		1	Desulfobacteraceae
		1	*alphaproteobacteria
		1	Pasteurellaceae
	DW	2	Bacteroidaceae
		1	Desulfobacteraceae
		4	Enterobacteriaceae ○
		1	Moraxellaceae
		2	Pseudomonadaceae
		1	Rhodospirillaceae
1		Idiomarinaceae ○	
1		Oceanospirillaceae ●	
1		Oxalobacteraceae	
4		Planococcaceae	
4	Shewanellaceae ○		
1	Spirochaetaceae		
Orne	UP	1	Oceanospirillaceae ●
		1	Enterobacteriaceae ○
	WW	1	Desulfomicrobiaceae
		1	Ignevibacteriaceae
		1	Oceanospirillaceae ●
		2	*Burkholderiales
	2	Campylobacteraceae	
	DW	1	Vibrionaceae ○
		1	*Burkholderiales
		1	Aeromonadaceae
2		Oceanospirillaceae ●	

38  
 39 Asterisks are for \*unclassified families. Open and full grey circles indicate SXT/R391 host families that  
 40 have been previously reported or also detected in this work when cloning and sequencing epicPCR  
 41 amplicons, respectively.

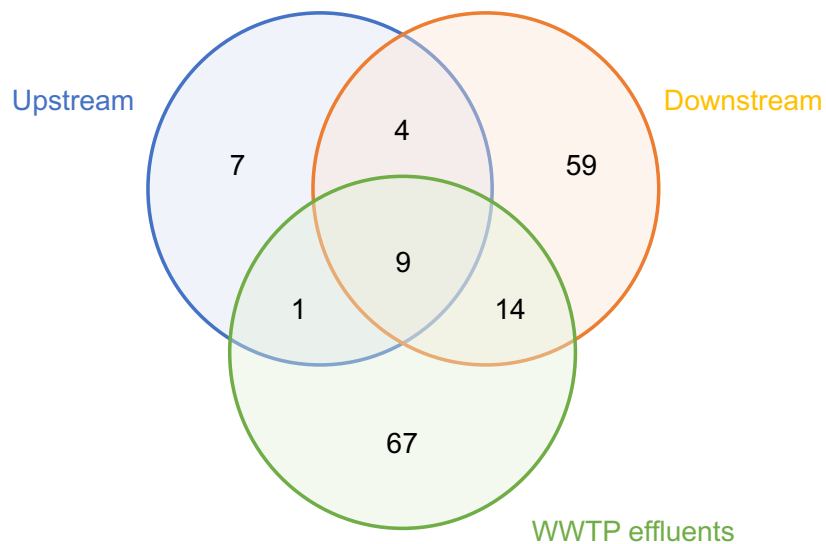
42  
 43



44

45 **Figure S1:** Electrophoretic analysis on agarose gel of the *traB* amplicons obtained after SYBR Green-  
 46 based qPCRs. The wells A to H indicate amplicons obtained using various environmental DNAs as  
 47 templates. C+ and C- wells indicate positive qPCR controls performed with DNA from strain *E. coli*  
 48 MG1656::SXT<sub>MO10</sub> and non-template qPCR controls, respectively. Similar results have been obtained  
 49 for qPCRs using a Taqman chemistry.

50



51

52

53 **Figure S2:** Venn diagram presenting the aggregated epicPCR results at the OTU level for the  
 54 experiments performed on the Orne River and the Meurthe River ecosystems. In downstream waters,  
 55 16.3% (14/86) of OTUs seem to exclusively originate from the WWTP effluents discharging in the rivers  
 56 whereas only 4.7% (4/86) of them would solely originate from the upstream waters.

57 **Additional references**

58 Götz, A., Pukall, R., Smit, E., Tietze, E., Prager, R., Tschäpe, H., ... Smalla, K., 1996. Detection and  
59 characterization of broad-host-range plasmids in environmental bacteria by PCR. *Applied and*  
60 *environmental microbiology*, 62(7), 2621-2628.

61 Hultman, J., Tamminen, M., Pärnänen, K., Cairns, J., Karkman, A., Virta, M., 2018. Host range of  
62 antibiotic resistance genes in wastewater treatment plant influent and effluent. *FEMS microbiology*  
63 *ecology*, 94(4), fiy038. <https://doi.org/10.1093/femsec/fiy038>

64 Spencer, S. J., Tamminen, M. V., Preheim, S. P., Guo, M. T., Briggs, A. W., Brito, I. L., ... & Alm, E. J.,  
65 2016. Massively parallel sequencing of single cells by epicPCR links functional genes with  
66 phylogenetic markers. *The ISME journal*, 10(2), 427-436. <https://doi.org/10.1038/ismej.2015.124>

67 Tremblay, J., Singh, K., Fern, A., Kirton, E. S., He, S., Woyke, T., ... Tringe, S. G., 2015. Primer and  
68 platform effects on 16S rRNA tag sequencing. *Frontiers in microbiology*, 6, 771.  
69 <https://doi.org/10.3389/fmicb.2015.00771>

