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.

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

Veronica L. Roman¹, Christophe Merlin¹, Sandrine Baron², Emeline Larvor², Laetitia Le Devendec², Marko P.J. Virta³, Xavier Bellanger^{1*}

¹ LCPME, CNRS, Université de Lorraine, Nancy, France

² Mycoplasmaology-Bacteriology and Antimicrobial Resistance Unit, Ploufragan-Plouzané-Niort Laboratory, French Agency for Food, Environmental and Occupational Health and Safety (Anses), Ploufragan, France

³ Department of Microbiology, University of Helsinki, Helsinki, Finland

* corresponding author:

LCPME - Equipe Microbiologie Environnementale

UMR 7564 CNRS - Université de Lorraine

Institut Jean Barriol

Campus Brabois Santé

Bâtiment AB, 3ème étage

9, Avenue de la Forêt de Haye - BP 20199

54505 Vandoeuvre-les-Nancy Cedex - FR

email: xavier.bellanger@univ-lorraine.fr

Abstract

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

Keywords

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

1 Introduction

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

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

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

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

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

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

2. Materials and Methods

2.1 Sample collection and bacteria libraries

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

collected on November 2017 and September 2019 in the Meurthe River ecosystem, and on September 2018, March 2019 and September 2019 in the Orne River ecosystem. The wastewater treatment plant (WWTP) located in Nancy processes around 12 000 m³ of wastewater a day from 120 000 residents, while the one located in Joeuf processes around 7 700 m³ of wastewater a day from 30 000 residents. Water was also collected from Mont-sur-Meurthe (MSM), Damelevières (DAM), Saint-Nicolas-de-Port (SNP), Art-sur-Meurthe (ASM) and Nancy (NAN) which are cities located along a 30 km long section of the Meurthe River, and from Joeuf (Upstream WWTP) and Beth barrage (Downstream WWTP) on the Orne River. Finally, this study also focused on the SXT/R391 content of 1794 bacteria isolated from four aquatic environments and wastewaters at the Anses Ploufragan-Plouzané laboratory. The 4 corresponding strain libraries mainly containing *V. cholerae* non-O1 non-O139, *Aeromonas spp.*, and *Pseudomonas spp.* strains collected over 19 years in France (Table S1). *Vibrio cholerae* and *Aeromonas spp.* strains were isolated and identified as described in Baron *et al.* (2013) and Baron *et al.* (2017), respectively. *Pseudomonas spp.* strains were isolated on RAPID[®] P.aeruginosa Agar (Bio-Rad) before being identified by PCR according to Benami *et al.* (2013). For epicPCR, 200 ml of water samples were centrifuged for 40 min at 8,000 g and 4°C. After discarding the supernatant, the remaining cell pellets were resuspended in 20% glycerol and stored at -20°C until use.

2.2 DNA extraction

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

2.3 Primers design and specificity checking

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

2.4 Quantitative PCR assays

SXT/R391 elements were detected and quantified by qPCR by targeting either the *traB* or the *int_{prfC}* genes from genomic and total environmental DNAs. The *traB* gene has been preferentially used as SXT/R391 ICEs exhibiting different *int* gene variants have been reported (Bioteau *et al.*, 2018). Quantitative PCRs were performed in triplicate using a “StepOnePlus Real-Time PCR System” (Applied Biosystems) with the “Taqman Universal Master Mix No UNG” when targeting the SXT/R391 ICEs, or with the “Power SYBR® Green PCR Master Mix” (Applied Biosystems) when the targeting class 1 integrons and 16S rRNA genes. SXT/R391 elements were quantified by qPCR according to the kit manufacturer, using primers at 40 µM (Table S2) and the following cycling conditions: initial denaturation 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. Class 1 integrons and eubacterial 16s rRNA genes were quantified according to De la Cruz Barrón *et al.* (2018). Plasmids pBELX2, containing a *traB* gene fragment from ICE SXT_{MO10}, or pNORM1, containing *intI1* and 16S rRNA gene fragments (Rocha *et al.*, 2020), linearized with either *NdeI* or *BamHI*, were used as qPCR standards for the quantification of SXT/R391 ICEs or class 1 integrons and environmental bacteria, respectively. These plasmids were extracted using the “Wizard plus SV Minipreps DNA Purification System” (Promega) according to the manufacturer recommendations. Their

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

2.5 16S rRNA gene metagenomics

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

2.6 EpicPCR

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

2.7 Sequencing data analysis

For recombinant plasmids, inserts were sequenced using the Sanger method (Genewiz) using primers listed in Table S2. The sequences obtained were analyzed using BLASTn (Altschul *et al.*, 1997). For metagenomics, 16S rDNA reads were joined with Pear (Zhang *et al.*, 2014) using default options, sequence quality was checked with FastQC (Andrews, 2020), primer sequences and short reads (< 330 bp) were removed using cutadapt (Martin, 2011). Unique sequences were identified using USEARCH -fastx_uniques command (Edgar, 2010). OTUs were clustered and reads were mapped to reference sequences with the USEARCH -cluster_otus command with the -minsize 2 parameter. Taxonomic classification of OTUs was done using the USEARCH -sintax command, using the ltp_16s_v123 database with the classifier cutoff set at 60. A representative sequence from each bacterial family (constituted of at least 10 reads) was aligned with MUSCLE (Madeire *et al.*, 2019), before a phylogenetic tree was built with Fasttree (Price *et al.*, 2009) and visualized in iTol (Letunic and Bork, 2019). Reads resulting from epicPCR amplifications were analyzed as described by Roman *et al.* (2021) and a confidence level (CL) in the detection of OTU determined, from level 1 (the lower), to level 5. This corresponds to the validation of an OTU based on its detection in epicPCR replicates as follows: (i) in only one replicate (CL=1), (ii) in either 2 out of 2 replicates of technical duplicates (so called a “2/2 technical replicates” validation method) or “2/3 technical” replicates (CL=2), (iii) in “3/3 technical” replicates (CL=3), (iv) in either “2/2 biological” or “2/3 biological” replicates (CL=4) or (v) “3/3 biological” replicates (CL=5).

3 Results and Discussion

3.1 Design of new set of SXT/R391 targeting primers

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

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

3.2 Abundance of STX/R391 ICEs in environmental bacteria libraries

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

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

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

3.3 Abundances of SXT/R391 ICEs in aquatic environments

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

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

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

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

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

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

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

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

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

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

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

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

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

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

3.5 Environmental significance of SXT/R391 host range broadness

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

that roughly 60% and 20% of the SXT/R391 ICEs described so far carry ARGs and metal/trace element resistance genes, respectively (Bioteau *et al.*, 2018). As example, ICEVchBan9 carries 6 ARGs (*sul2*, *floR*, *strAB*, *tetA*, *dfrA1*,) conferring resistance to 5 antibiotic classes, and 1 metal/trace element resistance gene (*zitB*). Only one fourth of these elements do not seem to encode cargo genes.

Otherwise, epicPCR allows very interesting and deep metagenomic analyses but can only offer an instant picture of a microbial community. This host range does not solely reflect the sets of different bacteria meet by the SXT/R391 ICEs, but may also have a deep mechanistic implication regarding their fate and behaviors in the context of their dissemination. First, the excision of these ICEs from the chromosome of their bacterial host, the earliest step of their transfer, has been shown as being either dependent or independent of host encoded protein factors. This dependence or independence is species-specific and may impair the ability some hosts to be good SXT/R391 donors (McLeod *et al.*, 2006). When present in two bacteria belonging to different species, the same SXT/R391 element can present basal transfer frequencies differing by more than one order of magnitude, and can exhibit transfer frequencies differently induce by a same concentration of an inducing molecule (Waldor *et al.*, 1996; Beaber *et al.*, 2004). Beyond these phenomena, SXT/R391 ICEs can carry a huge allelic variety of core genes and promoters (Bioteau *et al.*, 2018), that again multiplies the kinds of behavior they can have in a same species. The study of another ICE, Tn916, has shown that the same element can be very stable in a species, while highly unstable in another one. This instability has been linked to a very high excision frequency of the ICE and its subsequent lost during cell divisions, likely because of the misregulation of the promoters controlling the ICE recombination module (Scott *et al.*, 1988). Indeed, as the consensus sequences of bacterial promoters vary between species, a same gene can be differently expressed in two species because of the distinct recognition efficiencies of its promoter by the sigma factors of the bacteria where it has transferred (Haugen *et al.*, 2008). Finally, as others MGEs, SXT/R391 ICEs carry addiction modules, actually toxin-antitoxin systems, that promote the maintenance of the acquired elements (Wozniak and Waldor, 2009). Any mutation or misregulation in the expression of such an addiction module can lead in the ICE loss at a high frequency, or may impact on the viability of the bacterial cells carrying an MGE. All in all, these different phenomena draw a complex picture where the same SXT/R391 element can excise, transfer, integrate and maintained over cell divisions at different frequencies when it is harbored by hosts belonging to different bacterial species. Thus, if SXT/R391 elements are present in a variety of hosts within a microbial community, they likely behave

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

4 Conclusion

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

5 Acknowledgements

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

6 Funding sources

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

7 CRediT author statement

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

8 Competing Interests

None to declare

8 References

- Aarestrup, F. M., Seyfarth, A. M., Emborg, H. D., Pedersen, K., Hendriksen, R. S., Bager, F., 2001. Effect of abolishment of the use of antimicrobial agents for growth promotion on occurrence of antimicrobial resistance in fecal enterococci from food animals in Denmark. *Antimicrobial agents and chemotherapy*, 45(7), 2054-2059. <https://doi.org/10.1128/AAC.45.7.2054-2059.2001>
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., Lipman, D. J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research*, 25(17), 3389-3402. <https://doi.org/10.1093/nar/25.17.3389>
- Andersson, D. I., Hughes, D., 2014. Microbiological effects of sublethal levels of antibiotics. *Nature Reviews Microbiology*, 12(7), 465-478. <https://doi.org/10.1038/nrmicro3270>
- Andrews S. FastQC. Babraham Bioinformatics [cited on July 6, 2020]. Accessible on: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Aubertheau, E., Stalder, T., Mondamert, L., Ploy, M. C., Dagot, C., Labanowski, J., 2017. Impact of wastewater treatment plant discharge on the contamination of river biofilms by pharmaceuticals and antibiotic resistance. *Science of the Total Environment*, 579, 1387-1398. <https://doi.org/10.1016/j.scitotenv.2016.11.136>
- Badhai, J., Das, S. K., 2016. Characterization of three novel SXT/R391 integrating conjugative elements ICEM_{fulInd1a} and ICEM_{fulInd1b}, and ICEM_{prChn1} identified in the genomes of *Marinomonas fungiae* JCM 18476T and *Marinomonas profundimaris* strain D104. *Frontiers in microbiology*, 7, 1896. <https://doi.org/10.3389/fmicb.2016.01896>
- Balado, M., Lemos, M. L., Osorio C, R., 2013. Integrating conjugative elements of the SXT/R391 family from fish-isolated *Vibrios* encode restriction–modification systems that confer resistance to bacteriophages. *FEMS microbiology ecology*, 83(2), 457-467. <https://doi.org/10.1111/1574-6941.12007>

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.7735b392bdc749baf5812d2096d331e>

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 *Nitrincola alkalilacustris* sp. nov. and *Nitrincola schmidtii* sp. nov., alkaliphilic bacteria isolated from
606 soda pans, and emended description of the genus *Nitrincola*. 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 Coetsee, 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%2Fid1110.051014>

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

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 ICEAp/2 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/

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).

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 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.

SUPPLEMENTARY MATERIAL

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

Veronica L. Roman¹, Christophe Merlin¹, Sandrine Baron², Emeline Larvor², Laetitia Ledevendec², Marko P.J. Virta³, Xavier Bellanger^{1*}

¹LCPME, CNRS, Université de Lorraine, Nancy, France

²Mycoplasmology-Bacteriology Unit, Ploufragan-Plouzané Laboratory, French Agency for Food, Environmental and Occupational Health and Safety (Anses), Ploufragan, France

³Department of Microbiology, University of Helsinki, Helsinki, Finland

* corresponding author: xavier.bellanger@univ-lorraine.fr

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

Name	a	b	c	d	Sequence (5' - 3')	Targeted gene	Reference
traB _{SXT-1}	X			X	CTGCATTRCTGAGGATTCAA	<i>traB</i> (SXT/R391 elements)	This work
traB _{SXT-2}	X	X		X	CGGCGATCAAGAAACACTCT	<i>traB</i> (SXT/R391 elements)	This work
traB _{SXT-p}	X				TACCYAACCGATTYAGAGCK	<i>traB</i> probe (SXT/R391 elements)	This work
int _{SXT-1}	X				TGAAAAACAATTTTGCCCTTG	<i>int</i> (SXT/R391 elements)	This work
int _{SXT-2}	X				ATCAGTTAGCTGGCTCGATG	<i>int</i> (SXT/R391 elements)	This work
int _{SXT-p}	X				CWCGRTRCGCTTTTACTACG	<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				ATTACGCGGCTGCTGG	16S rRNA	Muyzer <i>et al.</i> , (1993)
traB _{SXT-1} 519R		X			GWATTACGCGGCKGCTGCTGCATTRCTG AGGATTCAA	<i>traB</i> epicPCR linker	This work
1492R		X			GGTTACCTTGTTACGACTT	16S rRNA	Lane (1991)
traB _{SXT-3} comp-inv		X			MGCTCTRAATCGGTTTGGTA	<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 GATCTtctGACTACHVGGGTATCTAATCC	16S rRNA reverse primer with Illumina TruSeq adapter	Hultman <i>et al.</i> , (2018)
Illum_785R_4		X			GTGACTGGAGTTCAGACGTGTGCTCTTCC GATCTctgagtGACTACHVGGGTATCTAATC C	16S rRNA reverse primer with Illumina TruSeq adapter	Hultman <i>et al.</i> , (2018)
SXT-3_comp-inv-Illumina		X			ATCTACACTCTTCCCTACACGACGCTCTT CCGATCTMGCTCTRAATCGGTTTGGTA	<i>traB</i> epicPCR nested primer with Illumina TruSeq adapter	This work
U519R-block10		X			TTTTTTTTTGWATTACGCGGCKGCTG/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	

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

Seq. type ^a	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 _{MO10} (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

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

^a: the reference sequence from ICE SXT_{MO10} (AY055428.1) (here, sequence type 3) is:
ctgcattactgaggattcaaaaggaagccattttacccaaccgatttagagcggacatcaaagagtgtttcttgatcgccg. The varying
nucleotides are presented as bold characters.

ctgcattgctgaggattcaaaaggaagccgtttacccaaccgatttagaactgatatacaaagagtgtttcttgatcgccg

Table S4: EpicPCR results and their associated level of confidence for the Meurthe and the Orne Rivers when reads from 16S sRNA gene sequencing were lower than 10 per OTU, and therefore not 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
	DW	2	Campylobacteraceae
		1	Vibrionaceae ○
		1	*Burkholderiales
		1	Aeromonadaceae
		2	Oceanospirillaceae ●

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

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

Figure S2: Venn diagram presenting the aggregated epicPCR results at the OTU level for the experiments performed on the Orne River and the Meurthe River ecosystems. In downstream waters, 16.3% (14/86) of OTUs seem to exclusively originate from the WWTP effluents discharging in the rivers 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), fyy038. <https://doi.org/10.1093/femsec/fyy038>
- 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>