Leibniz-Institut für Zoo- und Wildtierforschung, Abt. 3 Wildtierkrankheiten &
Leibniz-Institut für Gewässerökologie und Binnenfischerei, Abt. 3 Experimentelle Limnologie

---

“Urban wastewater and lakes as habitats for bacteria and potential vectors for pathogens”

(IRG III, LFV INFECTIONS’21)

Publikationsbasierte Dissertation

zur Erlangung des akademischen Grades

"doctor rerum naturalium"

(Dr. rer. nat.)

in der Wissenschaftsdisziplin "Mikrobiologie"

eingereicht an der

Mathematisch-Naturwissenschaftlichen Fakultät

der Universität Potsdam

von

Daniela Numberger

Potsdam, März 2019

Hauptbetreuer: Prof. Dr. habil.Hans-Peter Grossart

Zweitbetreuer: Prof. Alex D. Greenwood, PhD

Gutachter: Prof. Dr .Hans-Peter Grossart, Prof. Alex D. Greenwood, PhD, Prof. Dr. Ulrich Schaible

Ort und Tag der Disputation: Potsdam, 19.09.2019

Published online at the
Institutional Repository of the University of Potsdam:
https://doi.org/10.25932/publishup-43709
https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-437095

“Thousands have lived without **Love,**

not one without **Water**. “

Wystan Hugh Auden

This dissertation is based on the following three manuscripts:

[My own contributions are indicated for each publication in brackets.]

(1) Numberger D, Ganzert L, Zoccarato L, Mühldorfer K, Sauer S, Grossart HP, Greenwood AD. **Characterization of bacterial communities in wastewater with enhanced taxonomic resolution by full-length 16S rRNA sequencing**. (submitted and under review). [Sampling and sample processing, complete lab work except sequencing, partly data collection, data analysis, conceptualisation and writing of manuscript]

(2) Numberger D, Riedel T, McEwen G, Nübel U, Frentrup M, Schober I, Bunk B, Spröer C, Overmann J, Grossart HP, Greenwood AD. **Genomic analysis of three *Clostridioides difficile* isolates from urban water sources.** *Anaerobe* (2019). doi: https://doi.org/10.1016/j.anaerobe.2019.01.002. [Sampling and sample processing, most of the lab work, data collection and data analysis, conceptualisation and writing of manuscript]

(3) Numberger D, Dreier C, Vullioud C, Gabriel G, Greenwood AD, Grossart HP. **Recovery of influenza A viruses from lake water and sediments by experimental inoculation.** (submitted and under revision). [Sampling and sample processing, complete lab work, data collection, data analysis, conceptualisation and writing of manuscript]

Manuscripts that are in preparation but not part of the PhD thesis:

Numberger D, Zoccarato L, Woodhouse J, Ganzert L, Sauer S, Grossart HP, Greenwood AD. Big City Life – Rural versus urban aquatic bacterial communities (in prep.)

In addition to the manuscripts in this thesis, I have co-authored the following paper:

Beier S, Holtermann P, Numberger D, Schott T, Umlauf L, Jürgens K. A metatranscriptomics-based assessment of small-scale mixing of sulfidic and oxic waters on redoxcline prokaryotic communities. *Environmental Microbiology* (2018). doi:10.1111/1462-2920.14499

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

bp	base pairs
BSA	bovine serum albumin
cm	centimeter
DNA	deoxyribonucleic acid
ECE	embyronated chicken eggs
g	gramm
h	hours
IAV	influenza A virus
L	liter
kb	kilobases (1,000 bp)
m	meter
M	molar
MDCK	Madin-Darby Canine Kidney
MRSA	methicillin-resistant *Staphylococcus aureus*
µL	microliter
µm	micrometer
µM	micromolar
mg	milligram
mL	milliliter
min	minutes
nm	nanometer
OTU	operational taxonomic unit
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PFU	plaque forming unit
pM	picomolar
RT-qPCR	real-time quantitative polymerase chain reaction
RNA	ribonucleic acid
sec	seconds
SPF	specific pathogen free
SMRT	single molecule real time
spp.	species (plural)
WWTP	wastewater treatment plant

ABSTRACT

Water is essential to life and thus, an essential resource. However, freshwater resources are limited and their maintenance is crucial. Pollution with chemicals and pathogens through urbanization and a growing population impair the quality of freshwater. Furthermore, water can serve as vector for the transmission of pathogens resulting in water-borne illness.

The Interdisciplinary Research Group III – 'Water' of the Leibniz alliance project INFECTIONS'21 investigated water as a hub for pathogens focusing on *Clostridioides difficile* and avian influenza A viruses that may be shed into the water. Another aim of this study was to characterize the bacterial communities in a wastewater treatment plant (WWTP) of the capital Berlin, Germany to further assess potential health risks associated with wastewater management practices.

Bacterial communities of WWTP inflow and effluent differed significantly. The proportion of fecal/enteric bacteria was relatively low and OTUs related to potential enteric pathogens were largely removed from inflow to effluent. However, a health risk might exist as an increased relative abundance of potential pathogenic *Legionella* spp. such as *L. lytica* was observed. Three *Clostridioides difficile* isolates from wastewater inflow and an urban bathing lake in Berlin ('Weisser See') were obtained and sequenced. The two isolates from the wastewater did not carry toxin genes, whereas the isolate from the lake was positive for the toxin genes. All three isolates were closely related to human strains. This indicates a potential, but rather sporadic health risk. Avian influenza A viruses were detected in 38.8% of sediment samples by PCR, but virus isolation failed. An experiment with inoculated freshwater and sediment samples showed that virus isolation from sediment requires relatively high virus concentrations and worked much better in Madin-Darby Canine Kidney (MDCK) cell cultures than in embryonated chicken eggs, but low titre of influenza contamination in freshwater samples was sufficient to recover virus.

In conclusion, this work revealed potential health risks coming from bacterial groups with pathogenic potential such as *Legionella* spp. whose relative abundance is higher in the released effluent than in the inflow of the investigated WWTP. It further indicates that water bodies such as wastewater and lake sediments can serve as reservoir and vector, even for non-typical water-borne or water-transmitted pathogens such as *C. difficile*.

ZUSAMMENFASSUNG

Wasser ist lebensnotwendig und somit eine essentielle Ressource. Jedoch sind unsere Süßwasser-Ressourcen begrenzt und ihre Erhaltung daher besonders wichtig. Verschmutzungen mit Chemikalien und Krankheitserregern, die mit einer wachsenden Bevölkerung und Urbanisierung einhergehen, verschlechtern die Qualität unseres Süßwassers. Außerdem kann Wasser als Übertragungsvektor für Krankheitserreger dienen und daher wasserbürtige Krankheiten verursachen.

Der Leibniz-Forschungsverbund INFECTIONS'21 untersuchte innerhalb der interdisziplinären Forschungsgruppe III – „Wasser", Gewässer als zentralen Mittelpunkt für Krankheiterreger. Dabei konzentrierte man sich auf *Clostridioides difficile* sowie aviäre Influenza A-Viren, von denen angenommen wird, dass sie in die Gewässer ausgeschieden werden. Ein weiteres Ziel bestand darin, die bakterielle Gemeinschaften eines Klärwerkes der deutschen Hauptstadt Berlin zu charakterisieren, um anschließend eine Bewertung des potentiellen Gesundheitsrisikos geben zu können.

Bakterielle Gemeinschaften des Roh- und Klarwassers aus dem Klärwerk unterschieden sich signifikant voneinander. Der Anteil an Darm-/Fäkalbakterien war relativ niedrig und potentielle Darmpathogene wurden größtenteils aus dem Rohwasser entfernt. Ein potentielles Gesundheitsrisiko konnte allerdings von potentiell pathogenen Legionellen wie *L. lytica* festgestellt werden, deren relative Abundanz im Klarwasser höher war als im Rohwasser. Es wurden außerdem drei *C. difficile*-Isolate aus den Klärwerk-Rohwasser und einem städtischen Badesee in Berlin (Weisser See) gewonnen und sequenziert. Die beiden Isolate aus dem Klärwerk tragen keine Toxin-Gene, wohingegen das Isolat aus dem See Toxin-Gene besitzt. Alle drei Isolate sind sehr nah mit humanen Stämmen verwandt. Dies deutet auf ein potentielles, wenn auch sporadisches Gesundheitsrisiko hin. (Aviäre) Influenza A-Viren wurden in 38.8% der untersuchten Sedimentproben mittels PCR detektiert, aber die Virusisolierung schlug fehl. Ein Experiment mit beimpften Wasser- und Sedimentproben zeigte, dass für die Isolierung aus Sedimentproben eine relativ hohe Viruskonzentration nötig ist. In Wasserproben ist jedoch ein niedriger Titer an Influenza A-Viren ausreichend, um eine Infektion auszulösen. Es konnte zudem auch festgestellt werden, dass sich „Madin-Darby

Canine Kidney (MDCK)“-Zellkulturen im Gegensatz zu embryonierten Hühnereiern besser eignen, um Influenza A-Viren aus Sediment zu isolieren.

Zusammenfassend lässt sich sagen, dass diese Arbeit mögliche Gesundheitsrisiken aufgedeckt hat, wie etwa durch Legionellen im untersuchten Berliner Klärwerk, deren relative Abundanz in geklärtem Abwasser höher ist als im Rohwasser. Desweiteren wird indiziert, dass Abwasser und Gewässer als Reservoir und Vektor für pathogene Organismen dienen können, selbst für nicht-typische Wasser-Pathogene wie *C. difficile*.

GENERAL INTRODUCTION

Water as both a crucial resource and environment for microorganisms

The Earth is covered by around 70% water which is essential to all living beings. However, fresh water makes up only about 2.5% of the total water on Earth. Therefore, the maintenance of clean fresh water resources is crucial. However it is threatened by scarcity and pollution, which can result in animal and human health risks. Fresh water resources are influenced by climate change, population growth and migration, urbanization and industrial development [1–4]. Furthermore, freshwater ecosystems are under particular ecological pressure via increasing temperatures due to global climate change and anthropogenic activities leading to chemical pollution and introduction of (alien) invasive species with dramatic subsequent changes in biodiversity [2, 5–8].

Aquatic environments harbour a broad spectrum of microorganisms including bacteria, viruses and protozoa. Bacteria play a crucial and important role in aquatic systems and drive globally important biogeochemical cycles. They provide nutrients for primary production by remineralisation processes, use dissolved/particular organic matter (DOC/POC) as a carbon source and serve as food for protozoan and the higher trophic food chain (**Figure 1**) [9–12]. In addition to carbon, bacteria take part in sulphur [13, 14], nitrogen [15, 16], phosphorus [17–19] and other important element cycles by oxidation or reduction of compounds, which are then available for other metabolic processes.

Aquatic bacterial communities are complex and well adapted to their environment. However, disturbances by environmental stress such as invasive species [20–22], temperature change, different atmospheric CO_2 concentrations, precipitation and UV radiation [23–25] can change the composition of the bacterial community probably resulting in an alteration of ecosystem functions. Furthermore, disturbed microbial communities might enable the growth of harmful or pathogenic microbes [26–28] which further constitutes a health risk for humans and animals.

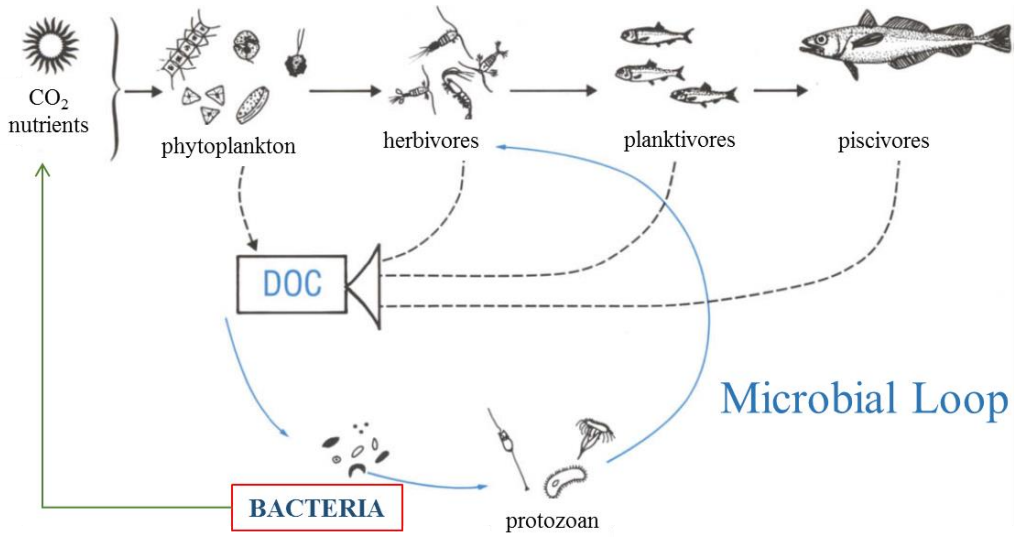


Figure 1: Simplified schematic representation of the food chain starting from primary production by phytoplankton to piscivores and the associated microbial loop. Bacteria use the dissolved organic matter (DOC) that is produced by the food chain and serve as food for protozoan (modified after Lalli and Parson, 1997 [9]).

Wastewater treatment plants

Wastewater treatment plants are crucial to maintain hygiene and to reduce pollution of the environment, particularly fresh water resources. Historically, urban drainage systems were constructed in ancient civilizations such as Indus, Minoan, Persians and Romans to avoid nuisance flooding, collect rainwater and transport waste [29–33]. Furthermore, wastewater, particularly waste from dry sewage systems, was used as fertilizer on farmland. The first modern-day wastewater treatment plants to reuse wastewater were constructed in the 1840's and 1850's [34, 35].

Cleaning of wastewater in a modern WWTP usually consists of a mechanical treatment followed by a biological treatment that can be combined with a chemical treatment. Most WWTPs have a disinfection step such as (partially) UV irradiation treatment or chlorination before releasing the effluent into the environment. The (activated) sludge that is produced during the treatment is dehydrated and incinerated (**Figure 2**). The mechanical treatment starts with screens to collect and remove coarse and fine solid material. Afterwards, sand and gravel are removed in grit chambers. Remaining solids initially settle in the primary clarifier, which are then part of the sludge. The core of the biological treatments consists of the aeration

tanks, which are rich in oxygen. Microorganisms remove nutrients such as phosphorus, ammonium, nitrate, nitrite and carbon and at the same time anaerobic (enteric) pathogens are eliminated. Nutrients can additionally be removed by chemical precipitation. In the secondary clarifier a final settling of solids can occur, before the water is disinfected with UV light or chlorine and released into the aquatic environment [36–38].

There are challenges associated with wastewater treatment, particularly in big cities. Although WWTPs provide cleaned wastewater, they can still be a source for pollution and contamination of the natural aquatic environment via WWTP effluent. WWTPs collect and concentrate wastewater from various sources including industry, hospitals and households. Thus, the wastewater can contain high levels of pharmaceuticals such as antibiotics, pathogens and (toxic) chemicals which might not be removed at all or not completely [39–41].

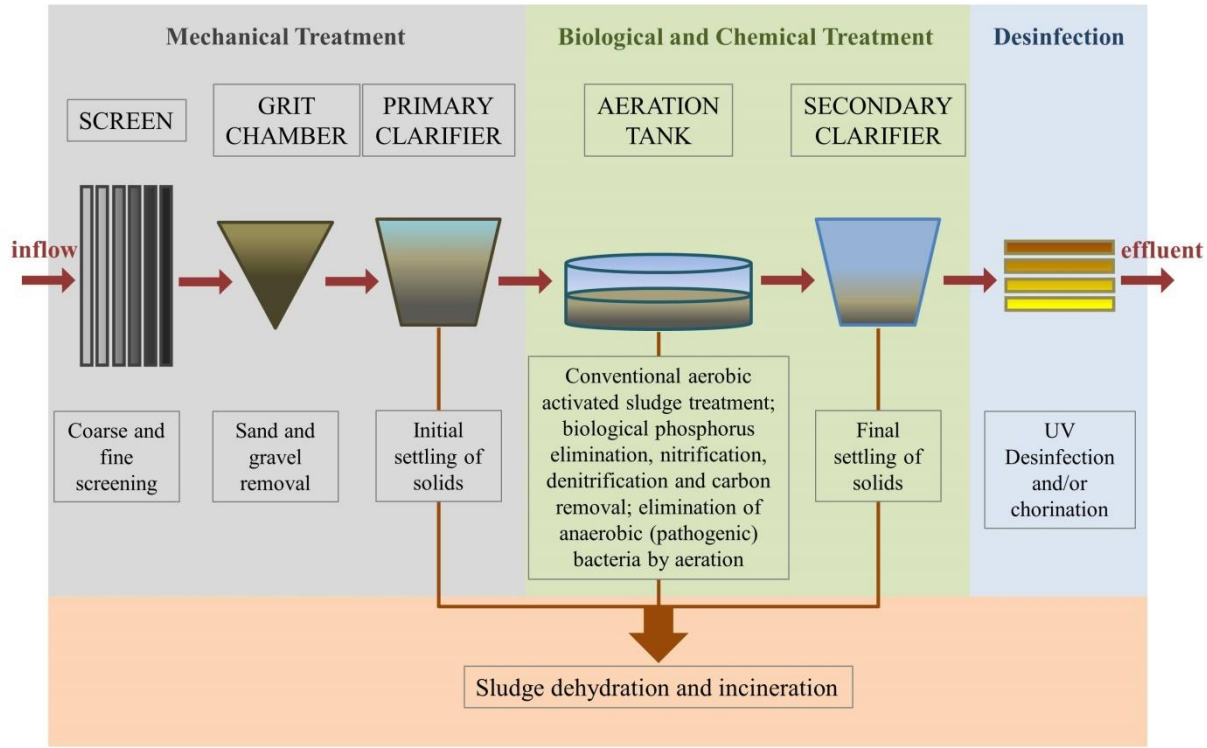


Figure 2: Schematic representation of the general wastewater treatment which consists of a mechanical treatment to remove sand, gravel and other solids, a biological treatment with aeration to remove anaerobic pathogens and most of the nutrients and a final disinfection step.

Pathogens in water and water as a vector

Water was considered first as a vector for pathogen transmission by the physician John Snow during the cholera outbreak in London, 1854. When he marked the locations of households where people died on a map, he could see that the cases were almost all restricted to the Golden Square area. This area differed from other areas in its source of drinking water which was a very polluted section of the Thames River [42–44]. Robert Koch found *Vibrio cholera* in the intestines and stool of cholera patients and identified *Vibrio cholera* as the causative agent for the disease [45, 46]. With this finding it was clear that fecal contamination of water is a serious health risk.

However, fecal contamination, particularly through (municipal or agricultural) wastewater is only one way of introducing pathogens into water sources [47]. **Figure 3** shows a schematic representation of possible ways pathogens can be introduced into natural open water sources. Wild animals, such as birds and rodents, can be a source of pathogens in aquatic environments. Waterfowl, for instance, can shed avian influenza A viruses via their feces [48], while rodents are known to be the natural reservoir for pathogenic leptospires that are shed via their urine [49–51]. Heavy rainfall and storm water can introduce pathogens from the environment such as *Clostridium perfringens*, *Salmonella* spp. and *Leptospira* spp. that are swept in from agricultural soil [52–58]. Humans and their recreational activities are also a possible source of pathogens [59, 60]. A study by Plano et al. [59], for instance, demonstrated that humans can shed methicillin-resistant *Staphylococcus aureus* (MRSA) into the water during swimming. Even bacteria of the genus *Enterococcus*, mainly fecal bacteria, can be released into the environment by bathers [60].

For many of the above mentioned pathogens such as MRSA, influenza A and *Leptospira* spp. it has been shown that they are persistent and stable in water and that they can remain infectious [61–64]. Typical cases of waterborne illness including possible symptoms and causative agents are listed in **Table 1**. There are many bacterial pathogens which are able to cause waterborne illness, e.g. *Aeromonas hydrophila, Campylobacter* spp., *Escherichia coli*, *Legionella* spp. *Leptospira* spp., *Pseudomonas aeruginosa*, *Salmonella* spp., *Vibrio cholera*, *Vibrio vulnificus* and *Mycobacterium marinum*.

However, there are also pathogens that have been frequently detected and isolated from the aquatic environment, but for which no reports of waterborne illness exist yet or they are quite rare [65–67]. Examples for such bacterial pathogens are MRSA and *Clostridioides difficile*. They are mainly known as nosocomial pathogens, but have been isolated from wastewater, freshwater and marine environments [66, 68–71]. For MRSA it has even been shown that it persist in water for at least 14 days [63]. This raises the question if they could become emerging waterborne pathogens in the future.

The concentration of pathogens, their survival and pathogenicity in water is influenced by a variety of environmental factors such as temperature, salinity, pH and nutrients [61–64]. Studies indicate that climate change, particularly ocean warming and heavy rain fall increase the incidence of water-borne or -transmitted disease, as many pathogens grow better at higher temperatures [72–74] as shown for *Vibrio* spp. in the Baltic Sea, for instance [75]. Moreover, there is evidence that aquatic pollution, particularly microplastics might serve as an additional vector for pathogens by being a substrate for forming biofilms [76, 77]. Urbanization and growing populations can have a negative effect on aquatic systems by increasing contamination and pollution [1, 56, 78–80]. Thus, urban waters might require more extensive monitoring programs in comparison to rural or undisturbed water.

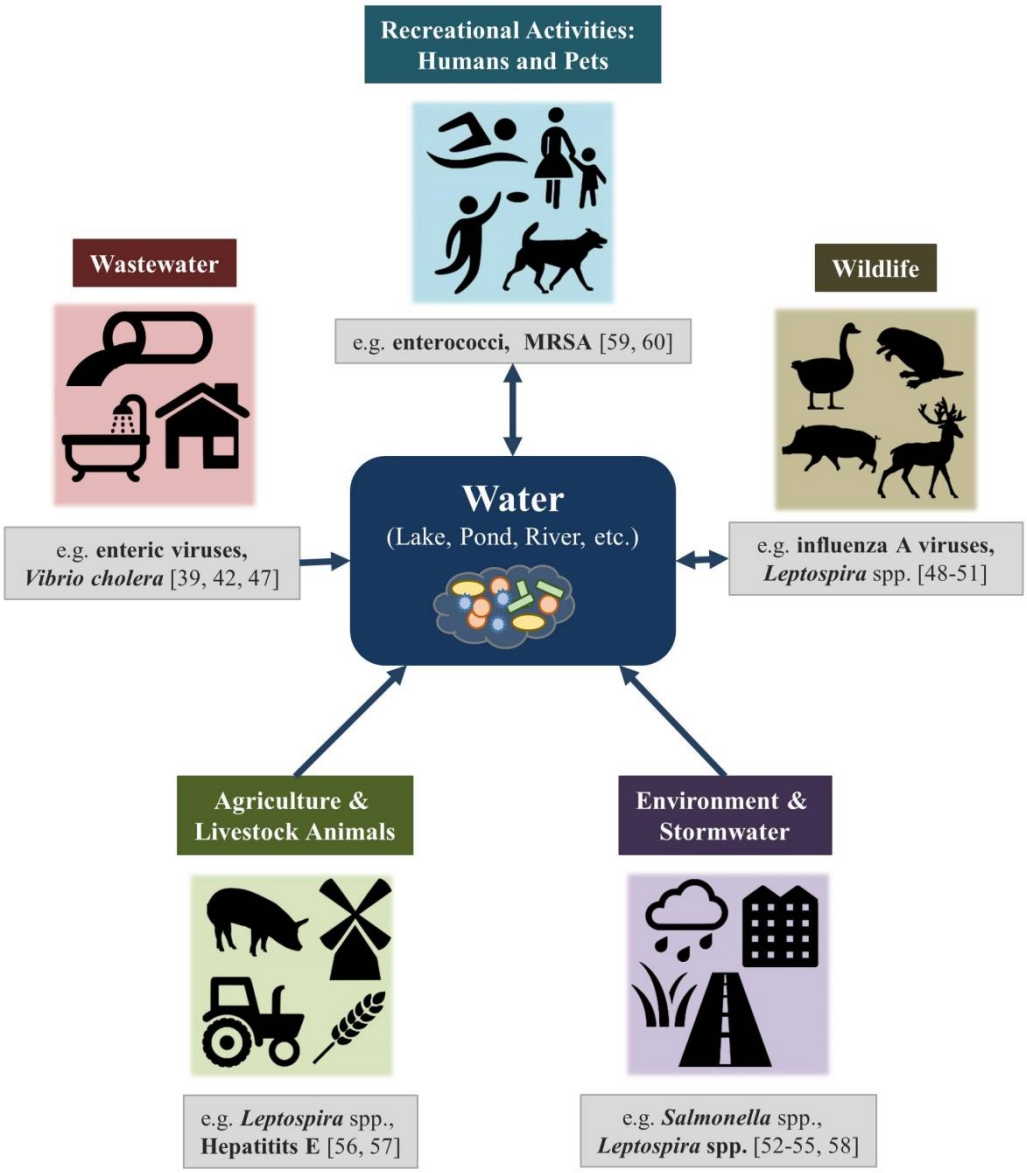


Figure 3: A schematic representation of potential ways, how pathogens can be introduced into natural open water sources. Wastewater, animals, the environment and storm water, humans and the agriculture can be sources of pathogens. The contaminated water can then serve as vector and infect animals and humans. Clipart black pictures are freely available from http://www.iconsdb.com.

Table 1: Typical waterborne diseases, their symptoms and causative agents (modified from Perkins et al. [81] including additional references for illness caused by *Salmonella* spp. [82, 83], *Campylobacter* spp. [84, 85] and *V. cholera* [42]).

Disease	Symptoms	Causative agents
Leptospirosis	Acute febrile illness,	*Leptospira* spp.
Legionnaires' disease, Pontiac Fever	pneumonia, influenza-like illness, fever, headache, muscle aches	*Legionella* spp.
Cryptosporidiosis	Diarrhea, abdominal pain, fever, nausea, vomiting, weight loss	*Cryptosporidium* spp.
Giardiasis	Abdominal cramps, arthralgias, diarrhea, hives, nausea, pruritus, vomiting	*Giardia intestinalis*
Other gastrointestinal infection	(bloody) diarrhea, fever, vomiting, gastroenteritis, hemolytic uremic syndrome in severe cases	Diarrheagenic or Shiga toxin-producing- *Escherichia coli*, *Salmonella* spp., *Campylobacter* spp.
Viral gastroenteritis	Abdominal pain, diarrhea, nausea, vomiting	Adenovirus, hepatitis A, norovirus, rotavirus
Cholera	Gastrointestinal symptoms	*Vibrio cholera*
Skin and soft tissue infections	Dermatitis, cellulitis, pruritus, skin erythema, soft tissue edema, skin necrosis, folliculitis	*Vibrio vulnificus*, *Aeromonas hydrophila*, *Mycobacterium marinum*, *Pseudomonas aeruginosa*
Algal bloom–related illness	Gastrointestinal symptoms, skin and lung irritation	Cytotoxins of algae (mainly cyanobacteria)

Thesis outline

This thesis project took place within the 'Interdisciplinary Research Group (IRG) III – Water' of the Leibniz Alliance Project 'INFECTIONS'21 - Transmission Control of Infections in the 21th Century' to examine whether water can act as a hub and/or reservoir for pathogens. We investigated the bacterial communities and the occurrence of selected pathogens in water and sediment samples. We analyzed five different lakes (Appendix **Figure A1**), and inflow and effluent of a wastewater treatment plant in Berlin, Germany and surroundings reflecting urban and rural habitats for comparison. In addition, a seasonal sampling and measurement of abiotic factors such as temperature, pH and nutrient concentration allowed us to analyse the data in an ecological context. The concept of the doctoral research is shown in **Figure 4** with following objectives:

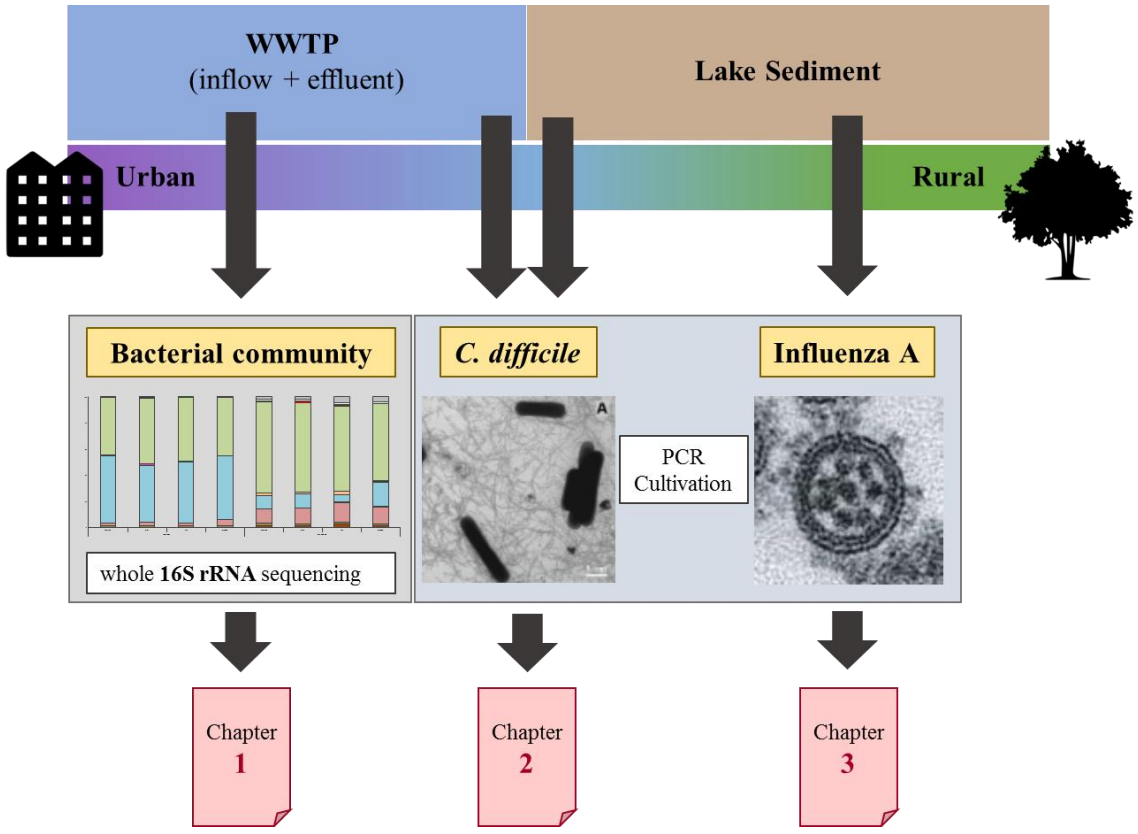


Figure 4: Concept of the doctoral research study within the IRG III – 'Water' of the Leibniz alliance project INFECTIONS'21. Sediment samples of lakes as well as inflow and effluent samples of a wastewater treatment plant (WWTP) were taken. Bacterial communities were defined by whole 16S rRNA gene sequencing in the WWTP samples (Chapter 1). Influenza A virus and *Clostridioides difficile* were selected as pathogens for further investigations (Chapters 2 and 3). Electron microscope images of influenza A virus **[86]** and *C. difficile* **[87]** are shown, respectively. Clipart black pictures are freely available from http://www.iconsdb.com.

1. Characterization of bacterial communities in wastewater with enhanced taxonomic resolution by full-length 16S rRNA sequencing (Chapter 1)

In big cities, WWTPs collect huge amounts of wastewater from different sources including households, hospitals and industry. Furthermore, WWTP can serve as habitat for numerous bacteria including pathogens as wastewater is rich in nutrients and particular carbon. Thus, there is a risk to WWTP employees of waterborne infection. With the effluent pathogens and other bacteria might reach natural water bodies and may represent a health risk for humans and animals. Whole 16S rRNA amplicon sequencing was performed with single molecule real time (SMRT) cell technology on the Sequel System of Pacific Biosciences (PacBio), USA to address following hypotheses:

H1: The bacterial communities differ from inflow to effluent indicating internal WWTP processes.

H2: With whole 16S rRNA sequences a better phylogenetic resolution (up to species level) can be achieved.

H3: WWTPs exhibit a potential risk for the contamination of the environment with pathogenic bacteria.

2. Genomic analysis of three *Clostridioides difficile* isolates from urban water sources (Chapter 2)

C. difficile, an anaerobic spore-forming bacterium, is mainly known as enteric, nosocomial pathogen producing toxins, with which they can cause diarrhoea, (pseudomembranous) colitis and other gastrointestinal disease, particularly in association with antibiotic treatment [88–92]. It has been shown that this pathogen can also be present and viable in the (aquatic) environment [68, 69, 93, 94]. Conventional and quantitative PCR (qPCR) of 16S rRNA, toxin A and toxin B genes combined with cultivation on selective chromID® *C. difficile* agar plates (bioMérieux, France) were performed. Genomes of isolates were sequenced with Illumina and PacBio technology. The hypothesis behind this study was:

H4: There is a higher prevalence of *C. difficile* occurrence in urban than in rural waters due to a higher risk/possibility of fecal contamination in the urban environment.

3. Recovery of influenza A viruses from lake water and sediments by experimental inoculation (Chapter 3)

Influenza A viruses cause disease in a variety of hosts such as humans, livestock and marine mammals [95–97]. The natural viral reservoir is birds, particularly waterfowl from which influenza A viruses are shed into water bodies via feces [98–102]. We applied conventional PCR of a 104-bp fragment of the matrix gene that is conserved through all influenza A viruses [103] followed by Sanger sequencing for confirmation. Further analyses included conventional PCR on hemagglutinin genes [104, 105] or the whole genome [106], as well as cultivation with Madin-Darby Canine Kidney (MDCK) cells and embryonated chicken eggs. With the assumption that sediment of lakes heavily used by waterfowl serve as reservoir for avian influenza A viruses, we tested following hypotheses:

H5: Influenza A virus frequently occur in sediment samples.

H6: Lake sediments serve as a reservoir for different strains of influenza A viruses.

H7: Influenza A viruses can be regrown from sediment samples and thus, maintain infectivity.

We additionally performed an experimental inoculation of sediment and freshwater samples with four different IAV strains. We aimed to define the minimum infectious dose that is required to recover IAV from those samples and to test the following hypothesis:

H8: The recovery efficiency of influenza A viruses from experimentally inoculated samples depends on the virus concentration, sample type (water vs. sediment), strain and the cultivation method used (MDCK cells vs. embryonated chicken eggs).

CHAPTER 1

Characterization of bacterial communities in wastewater with enhanced taxonomic resolution by full-length 16S rRNA sequencing

Characterization of bacterial communities in wastewater with enhanced taxonomic resolution by full-length 16S rRNA sequencing

Daniela Numberger[1], Lars Ganzert[2,3,4], Luca Zoccarato[2], Kristin Mühldorfer[1], Sascha Sauer[5], Hans-Peter Grossart[2,6,7], Alex D. Greenwood[1,8*]

1 Leibniz Institute for Zoo and Wildlife Research, Alfred-Kowalke-Straße 17, 10315 Berlin, Germany

2 Leibniz Institute of Freshwater Ecology and Inland Fisheries, Zur alten Fischerhütte 2, 16775 Stechlin, Germany

3 GFZ German Research Centre for Geosciences, Section 5.3 Geomicrobiology, Telegrafenberg C-422, 14473 Potsdam, Germany

4 University of Göttingen, Experimental Phycology and Culture Collection of Algae (SAG), Nikolausberger Weg 18, 37073 Göttingen, Germany

5 Max Delbrück Center for Molecular Medicine in the Helmholtz Association (BIMSB/BIH), Robert-Rössle-Strasse 10, 13092 Berlin, Germany

6 University of Potsdam, Institute of Biochemistry and Biology, Maulbeeralle 2, 14469 Potsdam, Germany

7 Berlin-Brandenburg Institute of Advanced Biodiversity Research (BBIB), Altensteinstrasse 32, 14195 Berlin, Germany

8 Freie Universität Berlin, Department of Veterinary Medicine, Institute for Virology, Robert von Ostertag-Straße 7-13, 14163 Berlin, Germany

* Corresponding author: Alex D. Greenwood, greenwood@izw-berlin.de

Running title: Characterization of bacterial communities in wastewater

Submitted and under review in Scientific Reports.

ABSTRACT

Wastewater treatment is crucial to environmental hygiene in urban environments. However, wastewater treatment plants (WWTPs) collect chemicals, organic matter, and microorganisms including pathogens and multi-resistant bacteria from various sources which may be potentially released into the environment via WWTP effluent. To better understand microbial dynamics in WWTPs, we characterized the bacterial community of the inflow and effluent of a WWTP in Berlin, Germany using full-length 16S rRNA gene sequences, which allowed for species level determination in many cases.

Significantly distinct bacterial communities were identified in the wastewater inflow and effluent samples. Dominant operational taxonomic units (OTUs) varied both temporally and spatially. Disease associated bacterial groups were efficiently reduced in their relative abundance from the effluent by the WWTP treatment process, except for *Legionella* and *Leptospira* species which increased from inflow to effluent.

This indicates that WWTP, while effective, may increase the release of some potentially pathogenic bacteria into the environment. The higher taxonomic resolution of full-length 16S rRNA genes allows for improved resolution of potential pathogenic taxa and other harmful bacteria which is required to reliably assess health risk.

Keywords: Wastewater treatment plant, bacterial community, full-length 16S rRNA gene sequencing, SMRT cell technology (PacBio®), urbanization

INTRODUCTION

Drinking water is a critical resource for which it is challenging to maintain hygiene in urban areas under persistent anthropogenic influence [1–4]. Pollutants and antibiotic resistant bacteria are constantly released by wastewater treatment plants (WWTP) into the environment which can result in human and (aquatic) animal health risk [5–11]. Treated sewage is also a major source of human-derived bacteria in the urban water environment, including potential pathogens that may survive the treatment process. Sewage inflow partly reflects the bacterial community of humans [12,13]. Accordingly, in two WWTPs in Hong Kong, China, pathogenic bacteria such as *Clostridium perfringens*, *Legionella pneumophila* and *Mycobacterium tuberculosis* like species were found to be common [14]. Thus, the WWTP effluent may not be completely depleted of (human) pathogens and the microbiome of the effluent and its nutrients may even promote the growth and proliferation of pathogenic bacteria in the environment. For example, Wakelin et al. [15] demonstrated that constant effluent input in combination with increased nutrient levels in the sediment downstream of a WWTP in Australia affected the bacterial community in the sediment substantially and increased the overall diversity. In addition, a study in rural Bangladesh revealed anthropogenic contamination of groundwater pumped from shallow tubewells with faecal bacteria from the genera *Shigella* and *Vibrio* [16] indicating the potential risk of faecal contamination of the natural environment via anthropogenic effluents.

WWTPs are considered hotspots for antibiotic resistant genes and for the spread of bacteria into the environment [9,17]. The presence of antibiotic resistant bacteria also increases the potential risk of gene transfer to non-resistant bacteria [18–20]. Several environmental bacteria are prone to developing multidrug resistance such as *Acinetobacter* spp., *Aeromonas* spp. and Pseudomonas spp. [21–23]. Adapted to humid and various aquatic environments these human-derived bacteria are part of the microbial communities in municipal WWTPs 24–26. For example, an increase of antibiotic resistant *Acinetobacter* spp. in WWTPs has been shown by Zhang et al. [21].

In addition, bacterial communities of wastewater include members of different taxonomic, biochemical and physiological groups, of which many provide advantageous functions for water cleaning, such as nutrient removal. However, these communities also contain bacteria

of human and animal origin which may interact with the bacterial communities of natural waters (e.g. rivers and lakes) in unpredictable ways. There are few studies comparing WWTP bacterial communities in inflow and effluent with the few undertaken restricted to a few countries, i.e. the USA, Hong Kong (China), and Spain [5,12,13,27,28]. Furthermore, these studies provide a relatively low taxonomic resolution since molecular identification is limited to short hypervariable regions of the 16S rRNA gene due to amplicon size constraints in sequencing on the Illumina or Roche 454 platforms [5,12,13,27,28]. Having only a restricted phylogenetic resolution, these methods do not allow for the reliable identification of human pathogenic bacteria in the environment. Full 16S rRNA gene sequencing, however, can provide improved taxonomic identification on the genus and species level. Therefore, we used single molecule real time (SMRT) sequening (PacBio® Sequel platform) to sequence full-length bacterial 16S rRNA gene PCR amplicons in order to improve the characterization of bacterial communities in wastewater samples. Inflow and effluent samples from a WWTP in Berlin (Germany) were collected every three months for one year to characterize in detail and compare the bacterial communities including potentially pathogenic bacteria.

RESULTS

Bacterial community composition and dominant OTUs

Using an OTU clustering cut-off of 99% sequence similarity, we were able to identify a total of 7,068 OTUs (initial data) of which 3,860 were left after rarefaction. Bacterial phyla and genera representing more than 1.0% of the total bacterial community were defined as dominant. Predominant phyla in the inflow were *Firmicutes*, *Proteobacteria*, *Bacteroidetes* and *Actinobacteria* with an average abundance of 52.2 ± 4.4%, 37.8 ± 4.7%, 4.9 ± 1.9%, and 2.2 ± 0.2%, respectively. In contrast, the effluent was dominated by *Proteobacteria* (54.8 ± 3.3%), *Bacteroidetes* (15.7 ± 1.1%), *Firmicutes* (14.3 ± 5,0%), *Actinobacteria* (2.6 ± 0.4%), *Planctomycetes* (2.9 ± 1.1%), *Acidobacteria* (1.3 ± 0.4%) and *Verrucomicrobia* (2.1 ± 0.4%) (**Fig. 1**, **Fig. 2**).

Families of the most dominant phyla *Proteobacteria*, *Bacteroidetes* and *Firmicutes* contributing in at least one sample more than 1% are shown in **Figure 4**. *Firmicutes* were dominated by *Acidaminococcaceae*, *Enterococcaceae*, *Eubacteriaceae*, *Lachnospiraceae*, *Peptostreptococcaceae*, *Ruminococcaceae*, *Streptococcaceae*, *Veillonellaceae*, *Christensenellacea* and *Clostridiaceae* 1 with *Lachnospiraceae* and *Ruminococcaceae* as most abundant. *Proteobacteria* were mainly represented by the families *Aeromonadaceae*, *Comamonadaceae*, *Enterobacteriaceae*, *Moraxellaceae*, *Neisseriaceae*, *Rhodobacteraceae*, *Rhodocyclaceae* *Campylobacteriaceae* and *Xanthomonadaceae*. For the phylum *Bacteroidetes* the families *Bacteroidaceae*, *Chitinophagaceae*, *Cytophagaceae*, *Flavobacteriaceae*, *Porphyromonadaceae*, *Prevotellaceae*, *Rikenellaceae* and *Saprospiraceae* were dominant.

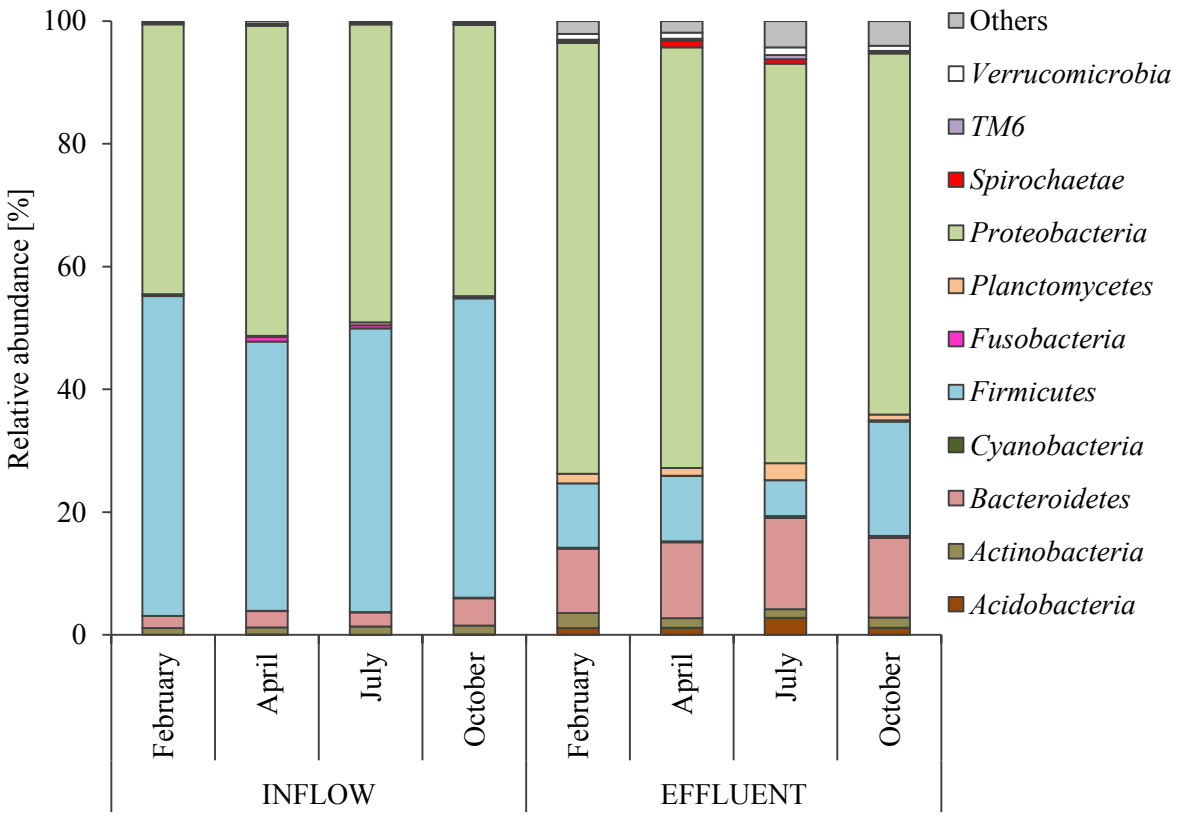


Figure 1: Composition of the bacterial community in inflow and effluent at the phylum level. The average abundance (after rarefaction) and only phyla contributing more than 1.0% to the total bacterial community are shown.

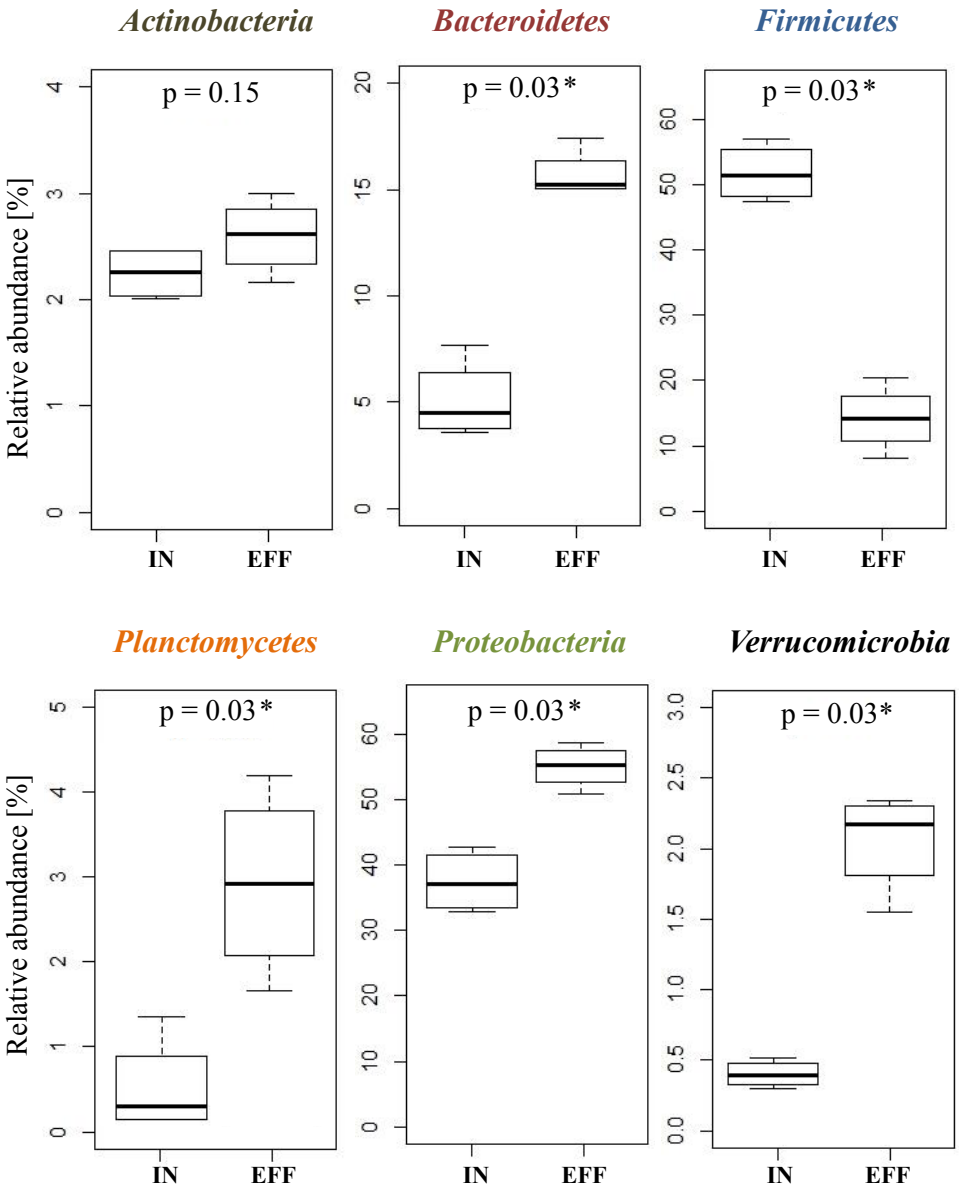


Figure 2: Boxplots showing the average relative abundance of the dominant bacterial phyla between inflow and effluent. The p-value (p) indicates the significance of the differences based on a PERMANOVA with $p < 0.05$ being significant.

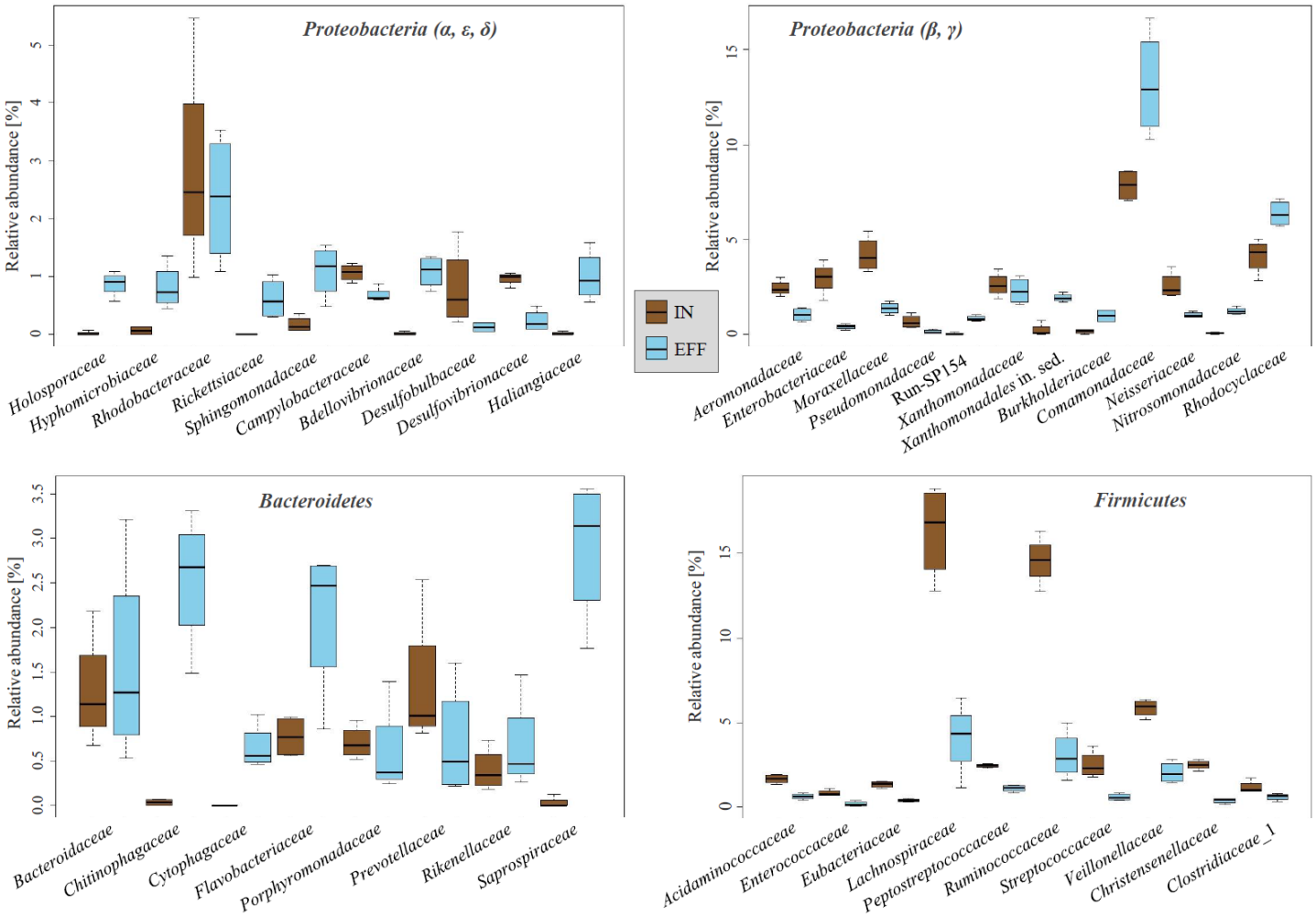


Figure 3: Average relative abundance of families contributing more than 1.0% to the total bacterial community in inflow and/or effluent. Only families of the three most abundant phyla *Proteobacteria*, *Bacteroidetes* and *Firmicutes* are shown.

OTUs contributing in at least one sample more than 1% to the total bacterial community are listed in **Table 1**. In addition, OTUs that significantly differ between inflow and effluent are marked in **Table 1** and belong to multiple genera: *Ca.* Accumulibacter, *Ca.* Competibacter, *Comamonadaceae* unclassifed, *Dechloromona*, *Nitrosomonas*, *Nitrospira*, *Paracoccus*, *Rhodoferax*, unclassified Run-SP154, *Simplicispira*, *Streptococcus*, unclassified *Saprospiraceae* and *Uruburuella*.

Using full-length 16S rRNA reads, we were able to reliably identify many OTUs at high taxonomic resolution (often at species level) by comparing them with reference sequences from known bacterial species (**Table 1**, **Table 2**) based on the global SILVA alignment (SINA Aligner) for rRNA genes 29. With the OTU sequences of the genus *Acinetobacter* we could show for example, that a better taxonomic resolution was produced in comparison with short-read sequences, which were generated by extracting a 477 bp fragment of the hypervariable regions V3-V4 according to the primer pair of Klindworth et al. [30]. Out of 113 OTUs related to the genus *Acinetobacter* 18 could be affiliated to a species when using full-length and 10 when using the hypervariable region V3-V4. However, the bootstrap values were always lower when using short sequences. Two OTUs yielded different phylogenetic results depending on sequence length (Supplementary **Table S1**, Supplementary **Fig. S9**).

Inflow versus effluent samples and dominant OTUs

Principal coordinate analysis defined two main clusters: inflow and effluent samples (**Fig. 4)**. Within the inflow cluster the samples from April and October were most similar to each other, whereas in the effluent cluster February and April or July and October samples were clustered more closely together. A permutational multivariate analysis of variance (PERMANOVA) revealed a significant difference between inflow and effluent samples with a p-value of 0.02. Furthermore, the dominant bacterial phyla *Bacteroidetes*, *Firmicutes*, *Planctomycetes* and *Verrucomicrobia* differed significantly in their relative abundance between inflow and effluent samples, whereas the phylum *Actinobacteria* did not show a significant difference between both sample groups (**Fig. 2**).

Table 1: Relative abundance (after rarefaction) of dominant OTUs (after rarefaction) with phylogenetic affiliation of inflow and effluent samples based on the global SILVA alignment (SINA Aligner) for rRNA genes [31]. Only OTUs contributing at least in one sample more than 1.0% are shown for the samples collected in February (Feb), April (Apr), July (Jul) and October (Oct). Coloured heat map indicates low abundance (green) and high abundance (red). OTUs that significantly change in abundance between inflow and effluent are indicated with red font colour (all are highly significant with p-values < 0.001). Arrows indicate an increase (↑) or decrease (↓) from inflow to effluent.

	INFLOW				EFFLUENT				Genus	Affiliated species	Sequence Identity
	Feb	Apr	Jul	Oct	Feb	Apr	Jul	Oct			
OTU000121	0,92	0,78	2,26	1,37	0,04	0,09	0,13	0,27	*Acetoanaerobium* ↓	-	
OTU000039	4,23	2,53	3,07	5,46	0,51	0,13	1,03	1,19	*Acidovorax* ↓	-	
OTU000062	4,32	3,00	1,66	2,35	0,47	0,36	0,18	0,78	*Acidovorax* ↓	*A. defluvii* (Y18616.1)	99.5%
OTU000055	6,58	3,27	2,60	2,53	0,25	0,98	0,67	0,63	*Acinetobacter* ↓	-	
OTU000065	2,39	4,43	1,30	1,63	0,34	0,40	0,63	0,76	*Aeromonas*↓	*A. media* (X60410.1)	97.8%
OTU000112	1,75	0,72	1,79	1,57	0,20	0,45	0,40	0,94	*Anaerosinus* ↓	*A. glycerini* (AJ010960.1)	98.5%
OTU000072	0,09	0,20	0,11	0,29	1,39	0,87	0,25	0,85	*Aquabacterium* ↑	-	
OTU000414	0,00	0,00	0,00	0,00	0,78	1,25	1,59	1,30	*Aquabacterium* ↑	-	
OTU000015	2,55	8,28	3,20	8,42	0,90	1,14	0,90	3,69	*Arcobacter* ↓	*A. cryaerophilus* (FR682113.1)	99.9%
OTU000307	0,45	0,07	0,34	0,51	0,18	0,40	0,11	1,10	*Bacteroides*	*B. graminisolvens* (AB547643.1)	99.8%
OTU000237	0,00	0,02	0,00	0,00	1,37	0,58	1,30	1,70	*Ca.* Accumulibacter ↑	-	
OTU000163	0,00	0,13	0,00	0,00	5,39	15,38	1,03	1,07	*Ca.* Competibacter ↑	-	
OTU000253	0,00	0,00	0,00	0,00	0,98	1,14	0,07	0,02	*Ca.* Nitrotoga ↑	*Candidatus* N. arctica (DQ839562.1)	99.5%
OTU001077	0,00	0,00	0,00	0,00	0,00	0,00	1,63	0,02	*Chryseobacterium* ↑	-	
OTU000006	0,00	0,00	0,00	0,00	3,04	1,19	0,16	1,95	*Comamonadaceae* uncl.↑	-	
OTU000040	0,00	0,07	0,00	0,00	6,20	2,26	0,11	0,78	*Comamonadaceae* uncl.↑	-	
OTU000071	0,00	0,00	0,00	0,00	2,01	1,05	1,16	0,74	*Comamonadaceae* uncl.↑	-	
OTU000188	0,65	0,87	1,01	0,96	0,00	0,02	0,02	0,09	*Comamonadaceae* uncl.↓	-	
OTU000585	0,00	0,00	0,00	0,00	1,12	0,22	0,16	0,04	*Comamonadaceae* uncl.↑	-	
OTU000031	4,90	2,75	6,11	5,55	0,43	0,27	0,16	1,54	*Comamonas* ↓	*C. denitrificans* (AF233880.1)	99.9%
OTU000377	0,00	0,00	0,00	0,00	0,00	0,00	2,89	1,39	*Cupriavidus*↑	-	

OTU000210	0,00	0,02	0,00	0,00	1,43	2,04	0,56	0,51	*Dechloromonas*↑	-	
OTU000187	1,45	1,14	0,34	1,05	0,07	0,07	0,02	0,11	*Enterococcus*↓	*E. aquimarinus* (EF204323.1)	99.7%
OTU000153	1,14	1,07	0,25	1,54	0,00	0,07	0,00	0,04	*Faecalibacterium*↓	*F. prausnitzii* (LQ500116.1)	99.7%
OTU000381	0,00	0,00	0,00	0,00	0,69	0,72	1,63	0,74	*Geothrix* ↑	*G. fermentans* (U41563.1)	97.8%
OTU000133	1,81	1,43	0,92	1,34	0,09	0,20	0,02	0,29	*Lachnospiraceae* uncl. ↓	-	
OTU000513	0,00	0,00	0,02	0,00	0,54	0,72	0,58	1,23	*Nitrosomonas* ↑	-	
OTU000519	0,00	0,00	0,00	0,00	0,22	0,13	1,19	1,05	*Nitrospira* ↑	-	
OTU000766	0,00	0,00	0,00	0,00	0,09	0,18	1,28	0,34	OM27 clade ↑	-	
OTU000467	0,34	0,60	1,75	0,09	0,02	0,00	0,02	0,00	*Paracoccus* ↓	*P. lutimaris* (KJ451483.1)	99.2%
OTU000209	0,22	0,45	0,25	0,81	1,05	0,40	0,60	1,19	*Peptostreptococcaceae* uncl.↑	-	
OTU000109	2,55	1,81	2,53	0,83	0,25	0,47	0,13	0,56	*Proteocatella* ↓	-	
OTU000466	0,00	0,00	0,00	0,00	4,23	0,72	0,02	0,16	*Rhodoferax* ↑	-	
OTU000339	0,00	0,02	0,00	0,00	0,45	0,27	1,23	1,59	Run-SP154 uncl. ↑	-	
OTU000756	0,00	0,00	0,00	0,00	2,13	0,78	0,47	0,13	*Simplicispira* ↑	*S. limi* (LC177120.1)	98.1%
OTU000123	0,51	0,96	5,35	0,78	0,07	0,04	0,09	0,07	*Streptococcus* ↓	*S. parasuis* (AB936273.1)	97.7%
OTU000387	0,11	0,11	2,22	0,27	0,04	0,04	0,02	0,04	*Streptococcus*↓	-	
OTU000685	0,04	0,04	1,48	0,11	0,00	0,00	0,00	0,00	*Streptococcus*↓	-	
OTU000873	0,02	0,00	1,12	0,00	0,02	0,00	0,02	0,02	*Streptococcus*↓	-	
OTU000128	1,84	1,50	0,74	1,10	0,38	0,47	0,11	0,47	*Subdoligranulum* ↓	-	
OTU000165	1,16	1,05	0,45	0,96	0,22	0,18	0,09	0,47	*Subdoligranulum* ↓	-	
OTU000078	0,27	1,28	1,84	1,03	0,92	1,16	4,07	1,88	*Thauera* ↑	-	
OTU000017	9,65	6,51	2,17	6,78	1,10	1,01	0,11	1,28	*Trichococcus*↓	*T. flocculiformis* (JF505981.1)	99.7%
OTU000098	0,00	0,04	0,00	0,00	0,78	0,85	1,19	0,67	*Saprospiraceae* uncl. ↑	-	
OTU001461	0,00	0,00	0,00	0,00	0,40	1,68	0,02	0,00	*Neisseriaceae* uncl. ↑	-	
OTU000099	2,57	2,71	3,02	0,94	0,13	0,07	0,02	0,07	*Uruburuella* ↓	*U. suis* (AJ586614.1)	99.9%
OTU000397	0,38	0,29	1,37	0,11	0,02	0,02	0,02	0,11	*Veillonella* ↓	-	
OTU000829	0,00	0,04	0,00	0,00	0,76	1,05	0,78	0,18	*Zoogloea* ↑	-	

Table 2: Affiliation of OTUs from potentially harmful bacterial genera and their presence in the inflow (IN) and effluent (EFF). This is based on the global SILVA alignment (SINA Aligner) for rRNA genes [31].

OTU	Presence	Genus	Affiliated species	Pairwaise sequence identity	Node support (bootstrap value)
OTU004330	IN, EFF	*Aeromonas*	*A. sharmana* (KC469704.1)	99.6%	97.3
OTU001375	IN, EFF		*A. sobria* (X60412.1)	99.0%	96.9
OTU027282	EFF		*A. jandaei* (X60413.1)	99.0%	87.4
OTU000640	IN, EFF		*A. australiensis* (HE611955.1)	99.2%	79.4
OTU000065	IN, EFF		*A. media* (X60410.1)	99.7%	100
OTU014486	IN	*Acinetobacter*	*A. beijerinckii* (KU308266.1)	99.7%	98.7
OTU017554	IN		*A. schindleri* (AJ278311.1)	97.9%	98.7
OTU025938	IN, EFF		*A. haemolyticus* (AY047216.1)	99.1%	77.5
OTU006694	IN		*A. celticus* (MBDL01000001.1)	99.7%	100
OTU004317	IN, EFF		*A. lwoffii* (KT369856.1)	98.6%	95.3
OTU012127	IN		*A. albensis* (KR611794.1)	98.7%	100
OTU004358	IN		*A. harbinensis* (KC843488.1)	98.6%	95.9
OTU011363	IN		*A. ursingii* (APQC01000001.1)	99.6%	100
OTU014283	IN		*A. oleivorans* (KF749341.1)	98.0%	78.0
OTU020140	IN		*A. radioresistens* (AM495259.1)	99.1%	100
OTU017557	IN		*A. rudis* (EF204258.3)	99.3%	100
OTU012742	IN		*A. baumannii* (X81660.1)	98.6%	100
OTU025639	IN		*A. indicus* (LC191521.1)	99.1%	88.2
OTU025642	IN		*A. gerneri* (APPN01000079.1)	99.4%	100
OTU020183	IN		*A. gandensis* (KM454858.1)	99.6%	100
OTU003228	IN, EFF		*A. junii* (KJ620866.1)	99.4%	100
OTU027605	IN	*Clostridium*	*C. frigidicarnis* (AF069742.1)	98.2%	100
OTU004478	IN, EFF		*C. perfringens* (AB610566.1)	99.7%	100
OTU034705	EFF		*C. colicanis* (AJ420008.1)	98.3%	100
OTU019193	IN		*C. paraputrificum* (AB627079.1)	98.5%	91.1
OTU022098	IN, EFF		*C. disporicum* (DQ855943.1)	99.6%	94.1
OTU024338	IN		*C. botulinum* type F (X68171.1)	99.5%	85.0
OTU037193	IN		*C. butyricum* (KY203641.1)	99.0%	100
OTU005080	IN, EFF		*C. beijerinckii* (LC071788.1)	99.6%	87.9
OTU014508	IN, EFF		*C. puniceum* (X71857.1)	98.1%	81.2
OTU009104	IN, EFF	*Legionella*	*L. feeleii* (LBHK01000101.1)	99.4%	100
OTU020580	EFF		*L. lytica* (X66835.1)	99.3%	100
OTU014279	IN	*Leptospira*	*L. alstonii* (CP015217.1)	99.8%	96.1
OTU006502	EFF	*Pseudomonas*	*P. pohangensis* (DQ339144.1)	98.0%	97.8
OTU001532	IN, EFF		*P. pseudoalcaligenes* (AJ628163.1)	98.6%	100
OTU030991	IN, EFF		*P. guangdongensis* (LT629780.1)	99.6%	100
OTU001564	IN, EFF		*P. alcaligenes* (CP014784.1)	99.8%	100
OTU017978	IN		*P. aeruginosa* (DQ641680.1)	99.6%	100
OTU025592	EFF		*P. psychrotolerans* (AJ575816.1)	99.0%	100
OTU020427	IN		*P. kunmingensis* (JQ246444.1)	98.5%	100
OTU032410	IN		*P. monteilii* (AF064458.1)	98.8%	82.9
OTU001416	IN, EFF		*P. baetica* (FM201274.1)	99.6%	99.5
OTU002122	IN		*P. gessardii* (KJ589457.1)	98.8%	86.1
OTU000948	IN, EFF		*P. palleroniana* (FNUA01000001.1)	98.8%	76.2
OTU007640	IN	*Yersinia*	*Y. massiliensis (EF179119.1)*	99.3%	98.9
OTU036842	IN		*Y. frederiksenii (AF366379.1)*	99.3%	99.9
OTU036840	IN		*Y. enterocolitica (CHYV01000006.1)*	99.3%	100
OTU006891	IN		*Y. intermedia (JX429054.1)*	99.5%	94.8

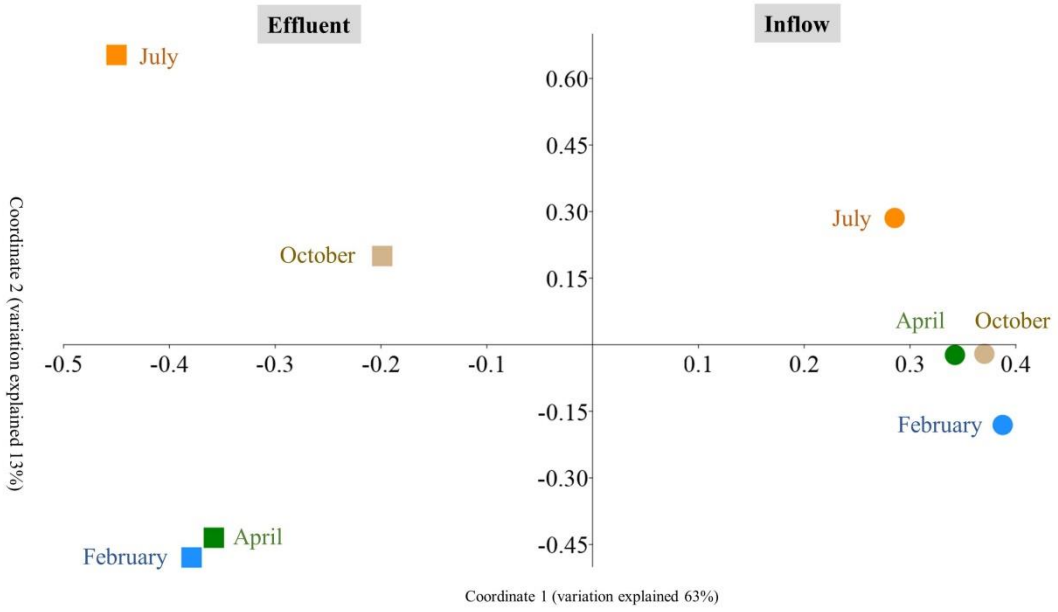


Figure 4: Principal coordinate analysis (PCoA) of the bacterial community based on Bray-Curtis similarity. Inflow and effluent samples are defined as circles and squares, respectively. Different sampling time points are indicated by blue colour for February, green colour for April, red colour for July and brown colour for October.

Phylogenetic analysis of genera that contain known pathogens

The advantage of full-length 16S rRNA gene sequencing was that, with some restrictions, more refined and reliable taxonomic assignment, even to the species level, was possible. While most of the previous studies used only the information of certain hypervariable regions of the 16S rRNA, we were able to use all phylogenetically relevant sites of the whole 16S rRNA gene. We therefore attempted to identify OTUs to a higher taxonomic level (e.g. species level), focusing on bacterial groups known to contain strains relevant for human health (**Table 2**). The analysis was carried out using maximum likelihood based phylogenetic approaches and including reference sequences from the SILVA database [31,32]. Three major groups of OTUs were identified representing (1) waterborne/-transmitted bacteria (i.e., *Legionella*, *Leptospira*, *Vibrio* and *Mycobacterium*) [33–36], (2) enteric bacteria (i.e., *Campylobacter*, *Clostridium*, *Salmonella*, *Shigella* and *Yersinia*) [37–42], and (3) environmental bacteria (i.e. *Acinetobacter*, *Aeromonas* and *Pseudomonas*) that include important nosocomial pathogens, which can also acquire multi-drug resistance [43–47].

Waterborne bacteria: Legionella, Leptospira, Mycobacterium and Vibrio

Legionella spp. and *Leptospira* spp. contributed to up to 0.9% and 1.0% to the bacterial community after rarefaction, respectively with increasing numbers from inflow to effluent (**Fig. 5**). Identified OTUs were closely related to *Legionella lytica*, *L. feeleii* (Supplementary **Fig. S1**), and *Leptospira alstonii* **(**Supplementary **Fig. S2)**. The genus *Mycobacterium* was only present in the October effluent samples with a relative abundance of 0.02%, whereas *Vibrio* was not detected in either the inflow or the effluent (**Fig. 5**).

Enteric bacteria: Campylobacter, Clostridium, Escherichia/Shigella, Salmonella and Yersinia

Campylobacter and *Salmonella* spp. were not detectable. The genus *Clostridium* (*sensu-stricto*) contributed between 0.1-0.9% to the bacterial community. *Escherichia*/*Shigella* and *Yersinia* decreased from inflow to effluent in relative abundance with *Yersinia* spp. being absent from the effluent samples (**Fig. 5**). According to our phylogenetic analyses probable species are *Clostridium perfringens*, *C. botulinum*, *C. butyricum* (Supplementary **Fig. S3**), *Yersinia massiliensis*, *Y. frederiksenii*, *Y. enterocolitica,* and *Y. media* (Supplementary **Fig. S4**). OTUs from the *Escherichia*/*Shigella* group did not show clear sequence similarity with any known species.

Environmental bacteria: Acinetobacter, Aeromonas and Pseudomonas

The genera *Acinetobacter, Aeromonas* and *Pseudomonas* were present in all samples, but their relative abundance decreased from inflow water to effluent in each of the sampled months (**Fig. 5**). *Acinetobacter* and *Aeromonas* spp. represented up to 9.5% and 5.8% of the bacterial community in the inflow, but only up to 1.3% and 1.1% in the effluent, respectively, while *Pseudomonas* spp. contributed only between 0.02% and 0.5% to the total bacterial community decreasing from inflow to effluent. OTUs were closely related to the described species *Acinetobacter beijerinckii*, *A. haemolyticus*, *A. baumannii* (Supplementary **Fig. S5**), *Aeromonas sharmana*, *A. media* (Supplementary **Fig. S6**), *Pseudomonas alcaligenes,* and *P. aeruginosa* (Supplementary **Fig. S7**).

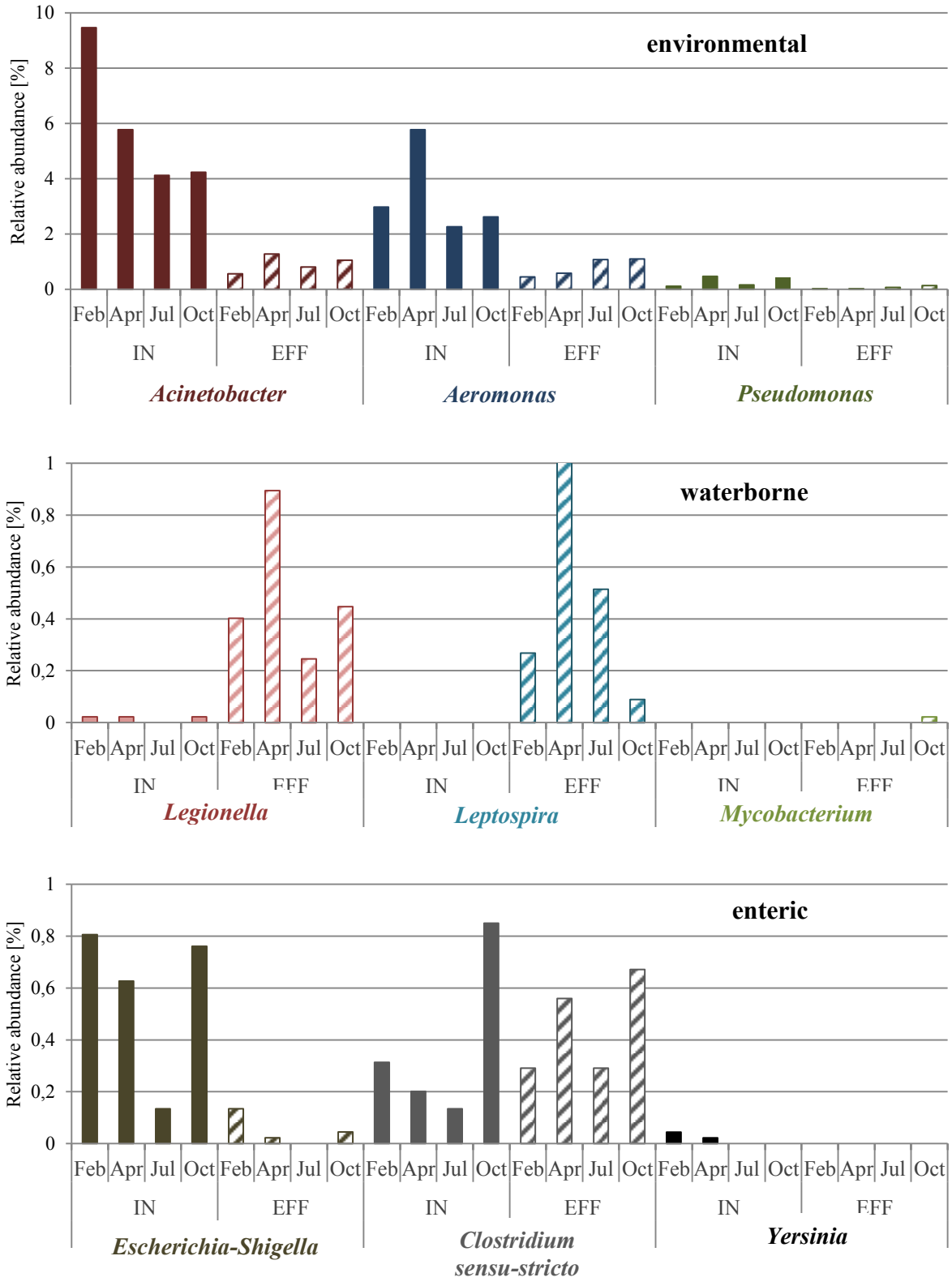


Figure 5: Relative abundance (after rarefaction) of genera with known potential pathogens. They were grouped in environmental, waterborne and enteric, and are shown for each sample of inflow and effluent.

DISCUSSION

Few studies describing the bacterial community in inflow water vs. effluent from a WWTP based on sequence data have been performed despite the potential for contamination of water bodies in highly urbanized areas [5,12,13,27,28]. Most studies have focused on specific bacterial groups or sampled only inflow water, activated sludge or the effluent. We found distinct compositional differences between the microbiomes of WWTP inflow water and effluent using a whole 16S rRNA gene sequencing approach.

At the phylum level there were two distinct clusters based on inflow and effluent specific bacterial communities, which showed only minor temporal differences (**Fig. 1**). Abiotic parameters such as oxygen concentration as well as competition among different bacterial species with different metabolic characteristics are very likely responsible for the observed differences in bacterial community composition in the WWTP inflow vs. the effluent. At the OTU level, however, there is evidence for seasonal or temporal differences (**Table 1**), but with only four time points sampled we could not draw any strong conclusions regarding seasonality.

While at the phylum level only minor differences occur between geographically distributed WWTPs, they differ strongly in the composition of the most abundant genera [5,12,27,28]. For example, our inflow samples shared seven dominant genera with the inflow water of a WWTP in Wisconsin (USA) [28] and nine [12] or three [27] genera with a WWTP in Hong Kong (China). The genera *Acinetobacter* and *Arcobacter* were dominant in all studies and are likely common members of WWTPs worldwide [5,12,27,28].

The differences could be further explained by other environmental parameters such as pH, temperature and salinity. The WWTP in Hong Kong, for example, treated wastewater has a salinity of 1.2% since it contains ca. 30% of seawater used for the toilet flushing system in Hong Kong [27]. This may possibly favour other bacterial groups in comparison to WWTPs that treat freshwater. Other reasons for the contrasting results might be the use of different small pore size filters for collecting bacteria and the application of different DNA extraction and sequencing methods. While part of the WWTP bacterial community reflects the human microbiome [13,48,49], some bacteria likely stem from industrial waste. Environmental bacteria

may reach the WWTP via rainfall and wildlife such as rodents inhabiting the drainage system. This might also explain observed regional differences in the bacterial community of WWTPs.

The dominant bacteria found in the current study can be useful or even necessary for the treatment process. *Comamonas denitrificans* has been shown to be a key organism in WWTPs and thus is very useful by its efficient denitrifying activity [50,51]. Its higher abundance in the outflow samples agreed with its presence in biofilms of the WWTP facility itself, including activated sludge [50–53]. Other species have been identified as abundant members in activated sludge and were suggested to be involved in nutrient removal including nitrite oxidation by *Nitrospira* spp. or enhanced biological phosphorus removal by *Simplicispira limi* [54–58], which were also abundant in the effluent of the current study.

Bacteria can be harmful for humans and animals by being pathogenic and/or by carrying antibiotic resistance genes. We grouped bacterial genera that contain known pathogenic species into three categories: waterborne, enteric or environmental bacteria that are prone to multidrug resistances.

Among waterborne bacteria *Vibrio cholera* is a well-studied waterborne pathogen [34,59] and has been found in WWTPs in Hong Kong, South Africa, USA and Brazil [14,60–62]. Contamination of WWTPs by cholera bacteria is likely human patient derived. As the incidence of cholera in Germany is negligible, this would explain why we never detected OTUs related to the genus *Vibrio*. *Legionella* and *Leptospira*, two other classical waterborne bacterial genera comprise known pathogenic species such as *Legionella pneumophila* and *Leptospira interrogans*. Interestingly, the relative abundance of OTUs belonging to these two genera increased from inflow to effluent samples indicating a potential health risk due to contamination of the environment or infection risk for WWTP workers. *Legionella* spp. are intracellular parasites and can replicate in free-living amoebae [63,64]. They likely form biofilms in the WWTP, which can promote bacterial growth and persistence in the aquatic environment [63,64]. In the current study, *L. lytica* and *L. feeleii* were identified as closest relatives (Supplementary **Fig. S1**). While the OTU related *to L. lytica* is exclusively present in the inflow samples, the OTU related to *L. feeleii* was detected in both inflow and effluent samples. Both species are known to cause pneumonia in humans when inhaled via aerosols [65–68] and may present a potential health risk as *Legionella* spp. in WWTP aerosols are not

unusual [69–71]. Wastewater, being enriched in nutrients and carbon, dissolved oxygen concentrations of 6.3-10.3 mg/L, and relatively high temperatures of 14.5-24.6 °C (Supplementary **Fig. S8**), provides favourable conditions for replication of *Legionella* spp. [72–74]. These pathogens remain challenging to control as they grow successfully within protozoa and biofilms, where they are relatively protected against disinfectants, grazers and other harsh environmental conditions [72].

The increase of *Leptospira* in the wastewater effluent could be associated with the presence of saprophytic leptospires that reproduce outside of a host and inhabit various aquatic environments [75,76]. Pathogenic *Leptospira*, however, can survive in water but do not reproduce outside of a host and thus may be introduced via an infected person or animals such as rodents, which are their natural reservoir and shed leptospires into their environment via urine [75,76]. Our phylogenetic analyses showed that the OTU affiliated with *L. alstonii* clustered with known pathogenic species such as *L. interrogans* and *L. mayottensis* [77,78] and was only present with one read in one of the inflow samples and thus, is likely derived from an infected human or rodent. All other *Leptospira* OTUs, were exclusively present in the effluent samples and belonged to saprophytic species such *as L. idonii* and *L. biflexa* [79,80] or were represented by their own cluster (Supplementary **Fig. S2**). This indicates that wastewater might favour the growth of saprophytic leptospires. While pathogenic leptopires grow much better at temperatures of around 30°C, saprophytic *Leptospira* spp. also replicate well at lower temperatures, as low as 10°C [81]. The temperatures of our wastewater samples varied between 14.5-24.6 °C during the sampled year (Supplementary **Fig. S8**). Furthermore, the ability to form biofilms may enhance their survival and/or replication in such an environment. However, as most of these OTUs seem to be related to saprophytic leptospires, we would assume a low health risk potential for humans and animals.

Enteric pathogens can secrete (entero-) toxins, which can damage the gastrointestinal tract of infected individuals [82–84]. They are part of the excreted faecal microbiota of humans in the WWTP inflow, but can also be introduced by animals such as rodents [85]. In the current study, *Clostridium* (*sensu-stricto*), *Escherichia-Shigella*, and *Yersinia* were mainly not abundant in the inflow, having a maximum relative abundance of 0.9% and were reduced in or absent from the effluent (**Fig. 5**). *Campylobacter* and *Salmonella* spp. were not detected at all, which

could mean that the sequencing depth was too low to detect them. While *Escherichia-Shigella* and *Yersinia* heavily decreased the relative abundance, *Clostridium* (*sensu-stricto*) remained mainly stable as observed previously [12]. These findings indicate that the wastewater treatment works well in removing enteric bacteria by introducing oxygen, preventing serious health risk.

Environmental bacteria such as *Acinetobacter*, *Aeromonas* and *Pseudomonas* spp. can be multidrug resistant [43,86,87] and some species also have a pathogenic potential such as *Pseudomonas aeruginosa* and *Acinetobacter baumannii* [88,89]. OTUs related to species like *P. aeruginosa* and *A. baumannii* were not abundant and only present in the WWTP inflow suggesting that the treatment procedures are effective against these species. Although the overall relative abundance of these three genera was reduced, they were not completely removed during the treatment process.

Pathogens can be strongly diluted in wastewater samples and masked by other much more abundant bacteria. Thus, the presence of pathogens could be greatly underestimated when using 16S rRNA data only. For instance, in a previous study we could detect and isolate *C. difficile* from the same samples used in the current study and even detect the *C. difficile* toxin genes via quantitative real-time PCR, but the 16S rRNA dataset did not provide any evidence for the presence of *C. difficile* [90]. Therefore, there are clearly limits to high throughput sequencing studies that involve a PCR step in terms of favouring abundant taxa. However, the current study provided evidence for the presence of other potential pathogens such as *Acinetobacter baumannii*, *Clostridium perfringens*, *Legionella lytica*, *Pseudomonas aeruginosa* and *Yersinia enterocolitica* by having the information of the full-length 16S rRNA gene, which may indicate that they are much more abundant than *C. difficile*, although still rare in the 16S rRNA dataset. Thus, further studies including isolation and cultivation methods are necessary to further investigate the presence and diversity of pathogens, to test for infectivity and to assess a realistic health risk. Particularly, water-adapted pathogens such as within the genus *Legionella* or *Leptospira* potentially increase in WWTPs and hence should be of great interest for health risk assessment, WWTP operation and waste management.

MATERIAL AND METHODS

Sampling

Untreated raw inflow water and treated effluent (no contact to the environment) of a wastewater treatment plant (WWTP) in Berlin, Germany, were sampled four times in 2016 (February 11th, April 15th, July 27th and October 20th). The selected WWTP treated municipal wastewater with only a minor percentage of industrial wastewater. It contains a mechanical treatment followed by biological one which includes biological phosphate elimination in combination with nitrification and denitrification, and the production of activated sludge. The effluent undergoes UV sterilization before its release in the environment. The exact location of the sampled WWTP cannot be disclosed due to a confidentiality agreement with the WWTP operators. The water samples were filtered through 0.22 µm Sterivex® filters (EMD Millipore, Germany) connected to a peristaltic pump (EMD Millipore, Germany) to concentrate bacteria and subsequently stored at -20°C. From the inflow water 20-35 mL could be concentrated on one filter, while from the effluent it was possible to filtrate 175-500 mL. Temperature, pH and dissolved oxygen were measured in the inflow samples with a digital thermometer (Carl Roth, Germany), pH multimeter EC8 (OCS.tec GmbH & CO. KG, Germany), Pen type, IP 67 dissolved oxygen meter (PDO-519, Lutron Electronic Enterprise CO., Taiwan), respectively.

DNA extraction

DNA was extracted from 0.22 µm Sterivex filters using the QIAamp DNA mini kit (Qiagen, Germany) following the protocol for tissue with some modifications. Briefly, the filters were cut into pieces and put into a 2 mL tube. 0.2 µm zirconium glass beads and 360 µL of buffer ATL were added and vortexed for 5 min at 3,000 rpm in an Eppendorf MixMate® (Eppendorf, Germany). Proteinase K (>600 mAU/ml, 40 µL) was added and incubated at 57°C for 1 h. After centrifugation for 1 min at 11,000 rpm, the supernatant was transferred to a new 2 mL tube and extraction was performed following the manufacturer's protocol.

Amplification of full-length 16S rRNA genes

Primers 27F (5'-AGRGTTYGATYMTGGCTCAG-3') and 1492R (5'-RGYTACCTTGTTACGACTT-3') were used with symmetric barcodes designed by Pacific Biosciences® (USA) for each sample. PCRs for each sample were run in triplicate and carried out in a total volume of 25 µL containing 12.5 µL MyFi^{TM} Mix (Bioline, UK), 9.3 µL water, 0.7 µL of bovine serum albumin (20 mg/mL; New England Biolabs, USA), 0.75 µL of each primer (10 µM) and 1 µL of DNA. The cycling program was as follows: denaturation at 95°C for 3 min, 25 cycles of 95°C for 30 sec, 57°C for 30 sec and 72°C for 60 sec and a final elongation at 72°C for 3 min. The quality and concentration of the PCR products were determined using a 4200 TapeStation with D5000 tapes and reagents (Agilent Technologies, USA). Equimolar sample mixes were used for library preparation.

Library preparation and sequencing

After bead purification with Agencourt AMPure XP (Beckman Coulter, USA), sequencing libraries were built using the SMRTbell Template Prep Kit 1.0-SPv3 following the guidelines in the amplicon template protocol (Pacific Biosciences, USA). DNA damage repair, end-repair and ligation of hairpin adapters were performed according to the manufacturer's instruction. DNA template libraries were bound to the Sequel polymerase 2.0 using the Sequel Binding Kit 2.0 (Pacific Biosciences, USA). The data collection per sample was done in a single Sequel SMRT Cell 1M v2 with 600 min movie time on the Sequel system (Pacific Biosciences, USA). We used a 5 pM on-plate loading concentration using Diffusion Loading mode and the Sequel Sequencing Plate 2.0 (Pacific Biosciences, USA).

Sequence Analysis

Circular consensus sequences (CCS) for each multiplexed sample were generated with the SMRT Analysis Software (Pacific Biosciences, USA) and used for further downstream analyses. An average of 7 Gb total output per SMRT cell was obtained, with an average CCS read length of 17 kb. Mean amplicon lengths of 1500 bp were confirmed. For further sequence processing Mothur 1.37 was used [91]. All sequences containing ambiguous bases, homopolymer stretches of >8 and shorter than 1,400 bp were removed. Sequences were

aligned using the align.seqs command in combination with the Silva v128_SSURef database. Reads that could not be aligned were removed and the remaining sequences were preclustered at 1% difference to account for potential PCR errors and then checked for chimeras using UCHIME in de novo mode [92]. Classification was done using classify.seqs using the RDP classifier implemented in Mothur [93] and Silva v128_SSURef database [31,32]. Sequences classified as Chloroplast-Mitochondria-unknown-Archaea were removed from the dataset. Operational taxonomic unit (OTU) clustering was done with VSEARCH (dgc mode; [94]) as implemented in Mothur, using a 99% similarity cutoff to nearly represent one species per OTU. This cutoff was used to resolve relationships among closely related bacteria that would be masked when using a cutoff of 97%. Phylogenetic analyses were performed with the ARB software using the LTPs128_SSU tree [95] and the SILVA database for bacterial 16S rRNA genes [31,32].

Phylogenetic analyses and statistics

Maximum-likelihood phylogenies (PhyML) were built with Jukes Cantor as the substitution model including 1,000 bootstrap replicates by using Geneious® 9.0.5 [96]. Rarefaction, log standardization of OTU counts and statistical analyses were performed using R version 3.5.1. To compare full-length with short read sequences, we restricted the sequences affiliated with the genus Acinetobacter using 16S rRNA primers to a 464 bp amplicon covering the hypervariable regions 3-4 [30]. To test for differential abundance of OTUs in the inflow versus the effluent, we used the exact negative binomial test in combination with the quantile-adjusted conditional maximum likelihood estimation of dispersion of the R package edgeR [97]. This analysis was based on TMM (trimmed mean of M values, where M is the log-fold-change of each OTU) normalized abundance data [98]. The test basically performed a pairwise comparison of OTU relative abundances between the two sample groups and an OTU was considered to respond significantly when the Bonferroni-corrected p-value was below 0.01.

AVAILABILITY OF DATA AND MATERIAL

The dataset generated and analysed during the current study are available from the Sequence Read Archive (SRA) of NCBI (National Center for Biotechnology Information) under the BioProject ID PRJNA484334 and SRA accession SRP156296.

ACKNOWLEDGEMENTS

We thank Claudia Quedenau for technical assistance to produce sequencing data, Daniele Y. Sunaga-Franze for computational support as well as the wastewater treatment plant in Berlin, which allowed us to do the sampling. The IGB-MIBI group is thanked for their discussions and technical support.

CONFLICT OF INTEREST

The authors confirm that they have no conflicts of interest related to the content of this article.

AUTHOR CONTRIBUTIONS STATEMENT

The study was designed by DN, HPG and ADG. DN collected and performed most of the practical work as well as data analysis and manuscript writing. LG processed the sequence data, LZ contributed by testing for differential abundance, KM supported the study with important scientific discussion, and SS provided the PacBio® Sequel sequencing expertise and sequencing data. All authors commented on the manuscript, read and agreed with the final version of the manuscript.

FUNDING

This work was supported by a project grant (IRG 3 - Water) from the Leibniz Research Alliance "INFECTIONS'21 - Transmission Control of Infections in the 21st Century" as funded by the Leibniz Association, Germany (SAS-2015-FZB-LFV) and the Leibniz SAW project "MycoLink" (SAW-2014-IGB).

REFERENCES

1. D'Arcy, B. & Frost, A. The role of best management practices in alleviating water quality problems associated with diffuse pollution. *Sci. Total Environ.* **265**, 359–367 (2001).
2. McLellan, S. L., Fisher, J. C. & Newton, R. J. The microbiome of urban waters. *Int. Microbiol. Off. J. Span. Soc. Microbiol.* **18**, 141–149 (2015).
3. Niemczynowicz, J. Urban hydrology and water management – present and future challenges. *Urban Water* **1**, 1–14 (1999).
4. Wang, J., Da, L., Song, K. & Li, B.-L. Temporal variations of surface water quality in urban, suburban and rural areas during rapid urbanization in Shanghai, China. *Environ. Pollut.* **152**, 387–393 (2008).
5. Marti, E., Jofre, J. & Balcazar, J. L. Prevalence of antibiotic resistance genes and bacterial community composition in a river influenced by a wastewater treatment plant. *PLOS ONE* **8**, e78906 (2013).
6. Miao, X.-S., Bishay, F., Chen, M. & Metcalfe, C. D. Occurrence of antimicrobials in the final effluents of wastewater treatment plants in Canada. *Environ. Sci. Technol.* **38**, 3533–3541 (2004).
7. Michael, I. *et al.* Urban wastewater treatment plants as hotspots for the release of antibiotics in the environment: A review. *Water Res.* **47**, 957–995 (2013).
8. Munir, M., Wong, K. & Xagoraraki, I. Release of antibiotic resistant bacteria and genes in the effluent and biosolids of five wastewater utilities in Michigan. *Water Res.* **45**, 681–693 (2011).
9. Rizzo, L. *et al.* Urban wastewater treatment plants as hotspots for antibiotic resistant bacteria and genes spread into the environment: a review. *Sci. Total Environ.* **447**, 345–360 (2013).
10. Szczepanowski, R. *et al.* Detection of 140 clinically relevant antibiotic-resistance genes in the plasmid metagenome of wastewater treatment plant bacteria showing reduced susceptibility to selected antibiotics. *Microbiol. Read. Engl.* **155**, 2306–2319 (2009).
11. Vajda, A. M. *et al.* Demasculinization of male fish by wastewater treatment plant effluent. *Aquat. Toxicol.* **103**, 213–221 (2011).
12. Cai, L., Ju, F. & Zhang, T. Tracking human sewage microbiome in a municipal wastewater treatment plant. *Appl. Microbiol. Biotechnol.* **98**, 3317–3326 (2014).
13. Newton, R. J. *et al.* Sewage reflects the microbiomes of human populations. *mBio* **6**, e02574-14 (2015).
14. Cai, L. & Zhang, T. Detecting human bacterial pathogens in wastewater treatment plants by a high-throughput shotgun sequencing technique. *Environ. Sci. Technol.* **47**, 5433–5441 (2013).
15. Wakelin, S. A., Colloff, M. J. & Kookana, R. S. Effect of wastewater treatment plant effluent on microbial function and community structure in the sediment of a freshwater stream with variable seasonal flow. *Appl. Environ. Microbiol.* **74**, 2659–2668 (2008).
16. Ferguson, A. S. *et al.* Comparison of fecal indicators with pathogenic bacteria and rotavirus in groundwater. *Sci. Total Environ.* **431**, 314–322 (2012).
17. Vaz-Moreira, I., Nunes, O. C. & Manaia, C. M. Bacterial diversity and antibiotic resistance in water habitats: searching the links with the human microbiome. *FEMS Microbiol. Rev.* **38**, 761–778 (2014).
18. Bennett, P. M. Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic resistance genes in bacteria. *Br. J. Pharmacol.* **153**, S347–S357 (2008).
19. Roe, E., Jones, R. J. & Lowbury, E. J. L. Transfer of antibiotic resistance between *Pseudomonas aeruginosa*, *Escherichia coli*, and other Gram-negative bacilli in burns. *Lancet* 149–52 (1971).
20. Courvalin, P. Transfer of antibiotic resistance genes between gram-positive and gram-negative bacteria. *Antimicrob. Agents Chemother.* **38**, 1447–1451 (1994).
21. Zhang, Y., Marrs, C. F., Simon, C. & Xi, C. Wastewater treatment contributes to selective increase of antibiotic resistance among *Acinetobacter* spp. *Sci. Total Environ.* **407**, 3702–3706 (2009).
22. Bonomo, R. A. & Szabo, D. Mechanisms of multidrug resistance in *Acinetobacter* species and *Pseudomonas aeruginosa*. *Clin. Infect. Dis.* **43**, S49–S56 (2006).
23. Kaskhedikar, M. & Chhabra, D. Multiple drug resistance in *Aeromonas hydrophila* isolates of fish. *Food Microbiol* **28**, 157–168 (2010).

24. Warskow, A. L. & Juni, E. Nutritional requirements of *Acinetobacter* strains isolated from soil, water, and sewage. *J. Bacteriol.* **112**, 1014–1016 (1972).
25. Kooij, D. van der, Visser, A. & Hijnen, W. a. M. Growth of *Aeromonas hydrophila* at low concentrations of substrates added to tap water. *Appl. Environ. Microbiol.* **39**, 1198–1204 (1980).
26. Sautour, M., Mary, P., Chihib, N. E. & Hornez, J. P. The effects of temperature, water activity and pH on the growth of *Aeromonas hydrophila* and on its subsequent survival in microcosm water. *J. Appl. Microbiol.* **95**, 807–813
27. Ye, L. & Zhang, T. Bacterial communities in different sections of a municipal wastewater treatment plant revealed by 16S rDNA 454 pyrosequencing. *Appl. Microbiol. Biotechnol.* **97**, 2681–2690 (2013).
28. McLellan, S. L., Huse, S. M., Mueller-Spitz, S. R., Andreishcheva, E. N. & Sogin, M. L. Diversity and population structure of sewage-derived microorganisms in wastewater treatment plant influent. *Environ. Microbiol.* **12**, 378–392 (2010).
29. Pruesse, E., Peplies, J. & Glöckner, F. O. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* **28**, 1823–1829 (2012).
30. Klindworth, A. *et al.* Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **41**, e1–e1 (2013).
31. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013).
32. Yilmaz, P. *et al.* The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Res.* **42**, D643–D648 (2014).
33. Hunter, P. R., Colford, J. M., LeChevallier, M. W., Binder, S. & Berger, P. S. Waterborne diseases. *Emerg. Infect. Dis.* **7**, 544 (2001).
34. Colwell, R. R. Infectious disease and environment: cholera as a paradigm for waterborne disease. *Int. Microbiol.* **7**, 285–289 (2004).
35. Cameron, D. & Jones, I. G. John Snow, the Broad Street Pump and Modern Epidemiology. *Int. J. Epidemiol.* **12**, 393–396 (1983).
36. Cann, K. F., Thomas, D. R., Salmon, R. L., Wyn-Jones, A. P. & Kay, D. Extreme water-related weather events and waterborne disease. *Epidemiol. Infect.* **141**, 671–686 (2013).
37. Ketley, J. M. Pathogenesis of enteric infection by *Campylobacter*. *Microbiology* **143**, 5–21 (1997).
38. Bartlett, J. G. *Clostridium difficile*: History of its role as an enteric pathogen and the current state of knowledge about the organism. *Clin. Infect. Dis.* **18**, S265–S272 (1994).
39. Smedley, J. G., Fisher, D. J., Sayeed, S., Chakrabarti, G. & McClane, B. A. The enteric toxins of *Clostridium perfringens*. in *Reviews of Physiology, Biochemistry and Pharmacology* 183–204 (Springer Berlin Heidelberg, 2005). doi:10.1007/s10254-004-0036-2
40. Broz, P., Ohlson, M. B. & Monack, D. M. Innate immune response to *Salmonella typhimurium*, a model enteric pathogen. *Gut Microbes* **3**, 62–70 (2012).
41. Wren, B. W. The *Yersiniae* — a model genus to study the rapid evolution of bacterial pathogens. *Nat. Rev. Microbiol.* **1**, 55–64 (2003).
42. Formal, S. B., Hale, T. L. & Sansonetti, P. J. Invasive enteric pathogens. *Rev. Infect. Dis.* **5**, S702–S707 (1983).
43. Figueira, V., Vaz-Moreira, I., Silva, M. & Manaia, C. M. Diversity and antibiotic resistance of *Aeromonas* spp. in drinking and waste water treatment plants. *Water Res.* **45**, 5599–5611 (2011).
44. Khan, N. H. *et al.* Isolation of P*seudomonas aeruginosa* from open ocean and comparison with freshwater, clinical, and animal isolates. *Microb. Ecol.* **53**, 173–186 (2007).
45. Guardabassi, L., Dalsgaard, A. & Olsen, J. E. Phenotypic characterization and antibiotic resistance of *Acinetobacter* spp. isolated from aquatic sources. *J. Appl. Microbiol.* **87**, 659–667 (1999).
46. Holmes, P., Niccolls, L. M. & Sartory, D. P. The ecology of mesophilic *Aeromonas* in the aquatic environment. *Genus Aeromonas* 127–150 (1996).

47. PARK, R. W. A. A Study of certain heterotrophic polarly flagellate water bacteria: *Aeromonas*, *Pseudomonas* and *Comamonas*. *Microbiology* **27**, 121–133 (1962).
48. Fisher, J. C. *et al.* Comparison of sewage and animal fecal microbiomes by using oligotyping reveals potential human fecal indicators in multiple taxonomic groups. *Appl. Environ. Microbiol.* **81**, 7023–7033 (2015).
49. Huttenhower, C. *et al.* Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207 (2012).
50. Andersson, S., Dalhammar, G., Land, C. J. & Rajarao, G. K. Characterization of extracellular polymeric substances from denitrifying organism *Comamonas denitrificans*. *Appl. Microbiol. Biotechnol.* **82**, 535 (2009).
51. Gumaelius, L., Magnusson, G., Pettersson, B. & Dalhammar, G. *Comamonas denitrificans* sp. nov., an efficient denitrifying bacterium isolated from activated sludge. *Int. J. Syst. Evol. Microbiol.* **51**, 999–1006 (2001).
52. Andersson, S. Characterization of bacterial biofilms for wastewater treatment. *DIVA* (2009).
53. Andersson, S., Kuttuva Rajarao, G., Land, C. J. & Dalhammar, G. Biofilm formation and interactions of bacterial strains found in wastewater treatment systems. *FEMS Microbiol. Lett.* **283**, 83–90 (2008).
54. Alawi, M., Off, S., Kaya, M. & Spieck, E. Temperature influences the population structure of nitrite-oxidizing bacteria in activated sludge. *Environ. Microbiol. Rep.* **1**, 184–190 (2009).
55. Lemaire, R., Yuan, Z., Blackall, L. L. & Crocetti, G. R. Microbial distribution of *Accumulibacter* spp. and *Competibacter* spp. in aerobic granules from a lab-scale biological nutrient removal system. *Environ. Microbiol.* **10**, 354–363 (2008).
56. Lücker, S. *et al. Nitrotoga*-like bacteria are previously unrecognized key nitrite oxidizers in full-scale wastewater treatment plants. *ISME J.* **9**, 708–720 (2015).
57. You, S. J., Hsu, C. L. & Ouyang, C. F. Identification of the microbial diversity of wastewater nutrient removal processes using molecular biotechnology. *Biotechnol. Lett.* **24**, 1361–1366 (2002).
58. Lu, S. *et al. Simplicispira limi* sp. nov., isolated from activated sludge. *Int. J. Syst. Evol. Microbiol.* **57**, 31–34 (2007).
59. Swerdlow, D. L. *et al.* Waterborne transmission of epidemic cholera in Trujillo, Peru: lessons for a continent at risk. *The Lancet* **340**, 28–32 (1992).
60. Barrett, T. J. *et al.* Use of Moore swabs for isolating *Vibrio cholerae* from sewage. *J. Clin. Microbiol.* **11**, 385–388 (1980).
61. Levine, M. M. *et al.* The Pathogenicity of nonenterotoxigenic *Vibrio cholerae* serogroup 01 biotype EI Tor isolated from sewage water in Brazil. *J. Infect. Dis.* **145**, 296–299 (1982).
62. Okoh, A. I. & Igbinosa, E. O. Antibiotic susceptibility profiles of some *Vibrio* strains isolated from wastewater final effluents in a rural community of the Eastern Cape Province of South Africa. *BMC Microbiol.* **10**, 143 (2010).
63. Bartram, J., Chartier, Y., Lee, J. V., Pond, K. & Surman-Lee, S. *Legionella* and the Prevention of Legionellosis. (World Health Organization, 2007).
64. Declerck, P. *et al.* Replication of *Legionella pneumophila* in biofilms of water distribution pipes. *Microbiol. Res.* **164**, 593–603 (2009).
65. Lee, J., Caplivski, D., Wu, M. & Huprikar, S. Pneumonia due to *Legionella feeleii*: case report and review of the literature. *Transpl. Infect. Dis.* **11**, 337–340 (2009).
66. Marrie, T. J., Raoult, D., La Scola, B., Birtles, R. J. & de Carolis, E. *Legionella*-like and other amoebal pathogens as agents of community-acquired pneumonia. *Emerg. Infect. Dis.* **7**, 1026–1029 (2001).
67. McNally, C., Hackman, B., Fields, B. S. & Plouffe, J. F. Potential importance of *Legionella* species as etiologies in community acquired pneumonia (CAP)1. *Diagn. Microbiol. Infect. Dis.* **38**, 79–82 (2000).
68. Muder, R. R. & Victor, L. Y. Infection due to *Legionella* species other than *L. pneumophila*. *Clin. Infect. Dis.* **35**, 990–998 (2002).

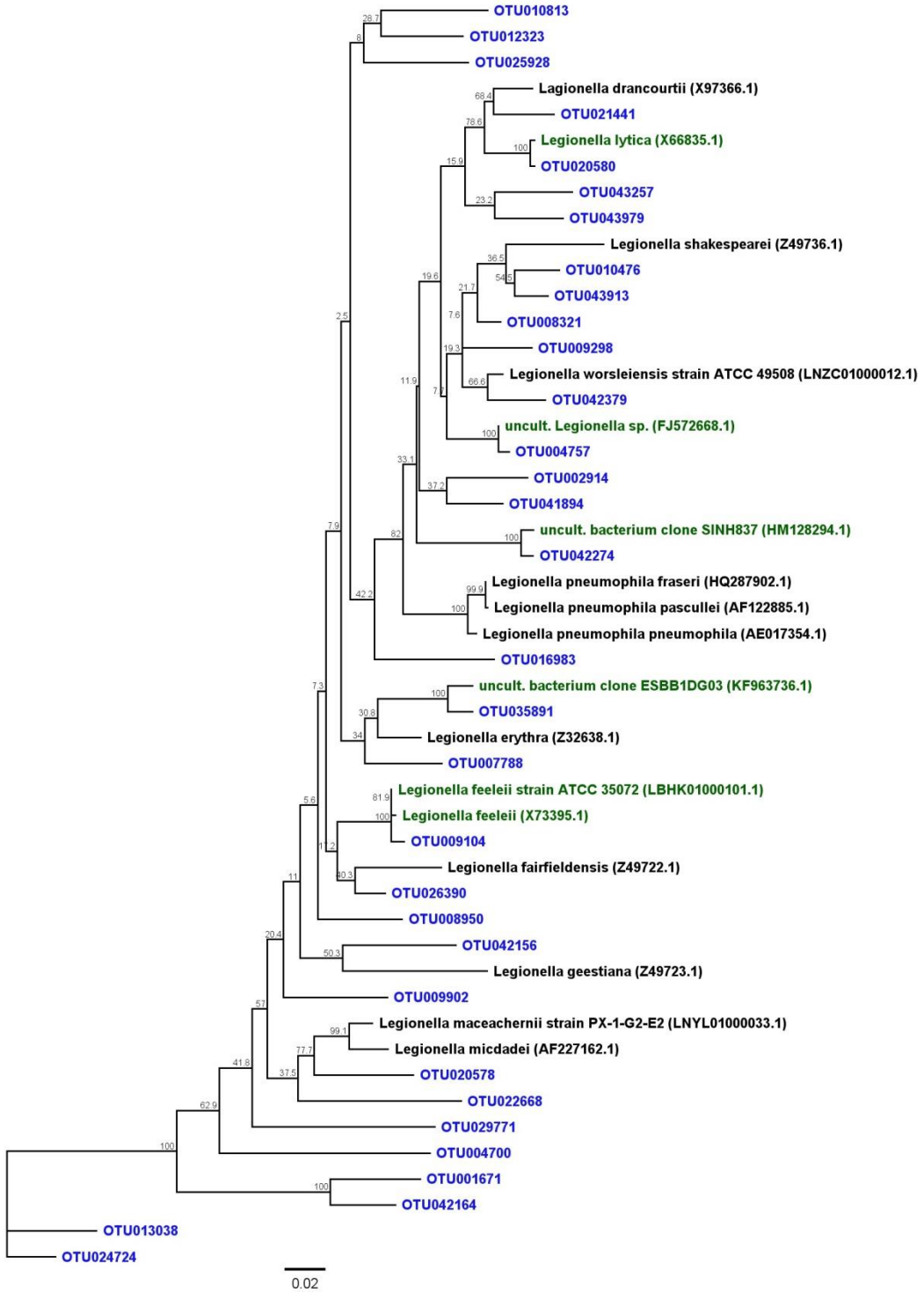
69. Medema, G., Wullings, B., Roeleveld, P. & Kooij, D. van der. Risk assessment of *Legionella* and enteric pathogens in sewage treatment works. *Water Sci. Technol. Water Supply* **4**, 125–132 (2004).
70. Pascual, L. *et al.* Bioaerosol emission from wastewater treatment plants. *Aerobiologia* **19**, 261–270 (2003).
71. Roll, B. M. & Fujioka, R. S. Detection of legionella bacteria in sewage by polymerase chain reaction and standard culture method. *Water Sci. Technol.* **31**, 409–416 (1995).
72. Molofsky, A. B. & Swanson, M. S. Differentiate to thrive: lessons from the *Legionella pneumophila* life cycle. *Mol. Microbiol.* **53**, 29–40 (2004).
73. Wullings, B. A. & Kooij, D. van der. Occurrence and genetic diversity of uncultured *Legionella* spp. in drinking water treated at temperatures below 15°C. *Appl. Environ. Microbiol.* **72**, 157–166 (2006).
74. Söderberg, M. A. & Cianciotto, N. P. Mediators of lipid A modification, RNA degradation, and central intermediary metabolism facilitate the growth of *Legionella pneumophila* at low temperatures. *Curr. Microbiol.* **60**, 59–65 (2010).
75. Aviat, F. *et al. Leptospira* exposure in the human environment in France: A survey in feral rodents and in fresh water. *Comp. Immunol. Microbiol. Infect. Dis.* **32**, 463–476 (2009).
76. Wynwood, S. J. *et al.* Leptospirosis from water sources. *Pathog. Glob. Health* **108**, 334–338 (2014).
77. Bourhy, P., Collet, L., Brisse, S. & Picardeau, M. *Leptospira mayottensis* sp. nov., a pathogenic species of the genus *Leptospira* isolated from humans. *Int. J. Syst. Evol. Microbiol.* **64**, 4061–4067 (2014).
78. Ren, S.-X. *et al.* Unique physiological and pathogenic features of *Leptospira interrogans* revealed by whole-genome sequencing. *Nature* **422**, 888–893 (2003).
79. Saito, M. *et al. Leptospira idonii* sp. nov., isolated from environmental water. *Int. J. Syst. Evol. Microbiol.* **63**, 2457–2462 (2013).
80. Picardeau, M. *et al.* Genome sequence of the saprophyte *Leptospira biflexa* provides insights into the evolution of *Leptospira* and the pathogenesis of leptospirosis. *PLOS ONE* **3**, e1607 (2008).
81. Johnson, R. C. & Harris, V. G. Differentiation of pathogenic and saprophytic leptospires I. growth at low temperatures. *J. Bacteriol.* **94**, 27–31 (1967).
82. Burgen, A. S. V., Dickens, F. & Zatman, L. J. The action of botulinum toxin on the neuro-muscular junction. *J. Physiol.* **109**, 10–24 (1949).
83. Delor, I. & Cornelis, G. R. Role of *Yersinia enterocolitica* Yst toxin in experimental infection of young rabbits. *Infect. Immun.* **60**, 4269–4277 (1992).
84. Wassenaar, T. M. Toxin production by *Campylobacter* spp. *Clin. Microbiol. Rev.* **10**, 466–476 (1997).
85. Arnone, R. D. & Walling, J. P. Waterborne pathogens in urban watersheds. *J. Water Health* **5**, 149–162 (2007).
86. van Dessel, H. *et al.* Identification of a new geographically widespread multiresistant *Acinetobacter baumannii* clone from European hospitals. *Res. Microbiol.* **155**, 105–112 (2004).
87. Odjadjare, E. E. *et al.* Prevalence of multiple antibiotics resistant (MAR) *Pseudomonas* species in the final effluents of three municipal wastewater treatment facilities in South Africa. *Int. J. Environ. Res. Public. Health* **9**, 2092–2107 (2012).
88. Mena, K. D. & Gerba, C. P. Risk assessment of *Pseudomonas aeruginosa* in water. *Reviews of Environmental Contamination and Toxicology Vol 201* 71–115 (Springer, Boston, MA, 2009). doi:10.1007/978-1-4419-0032-6_3
89. Peleg, A. Y., Seifert, H. & Paterson, D. L. *Acinetobacter baumannii*: Emergence of a successful pathogen. *Clin. Microbiol. Rev.* **21**, 538–582 (2008).
90. Numberger, D. *et al.* Genomic analysis of three *Clostridioides difficile* isolates from urban water sources. *Anaerobe* (2019). doi:10.1016/j.anaerobe.2019.01.002
91. Schloss, P. D. *et al.* Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* **75**, 7537–7541 (2009).
92. Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C. & Knight, R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**, 2194–2200 (2011).

93. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73**, 5261–5267 (2007).
94. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**, e2584 (2016).
95. Ludwig, W. *et al.* ARB: a software environment for sequence data. *Nucleic Acids Res.* **32**, 1363–1371 (2004).
96. Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).
97. Robinson, M. D. & Smyth, G. K. Small-sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics* **9**, 321–332 (2008).
98. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010)

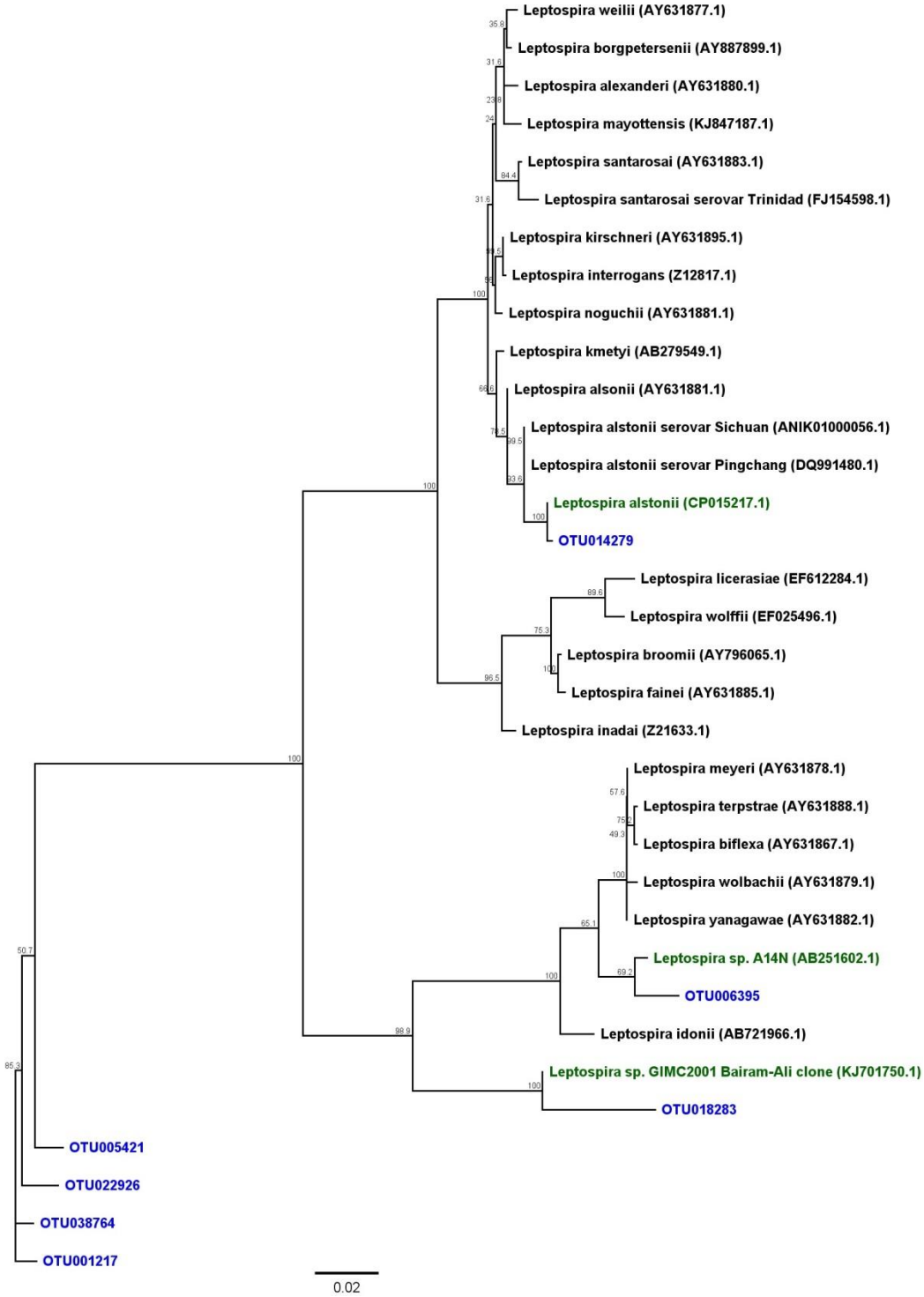
CHAPTER 1: SUPPLEMENTARY MATERIAL

Table S1: Comparison of the phylogenetic affiliation of *Acinetobacter*-OTUs between short (hypervariable region V3-V4) and full-length 16S rRNA sequence information. OTUs which could only be affiliated to a species when having full-length information are highlighted in bold.

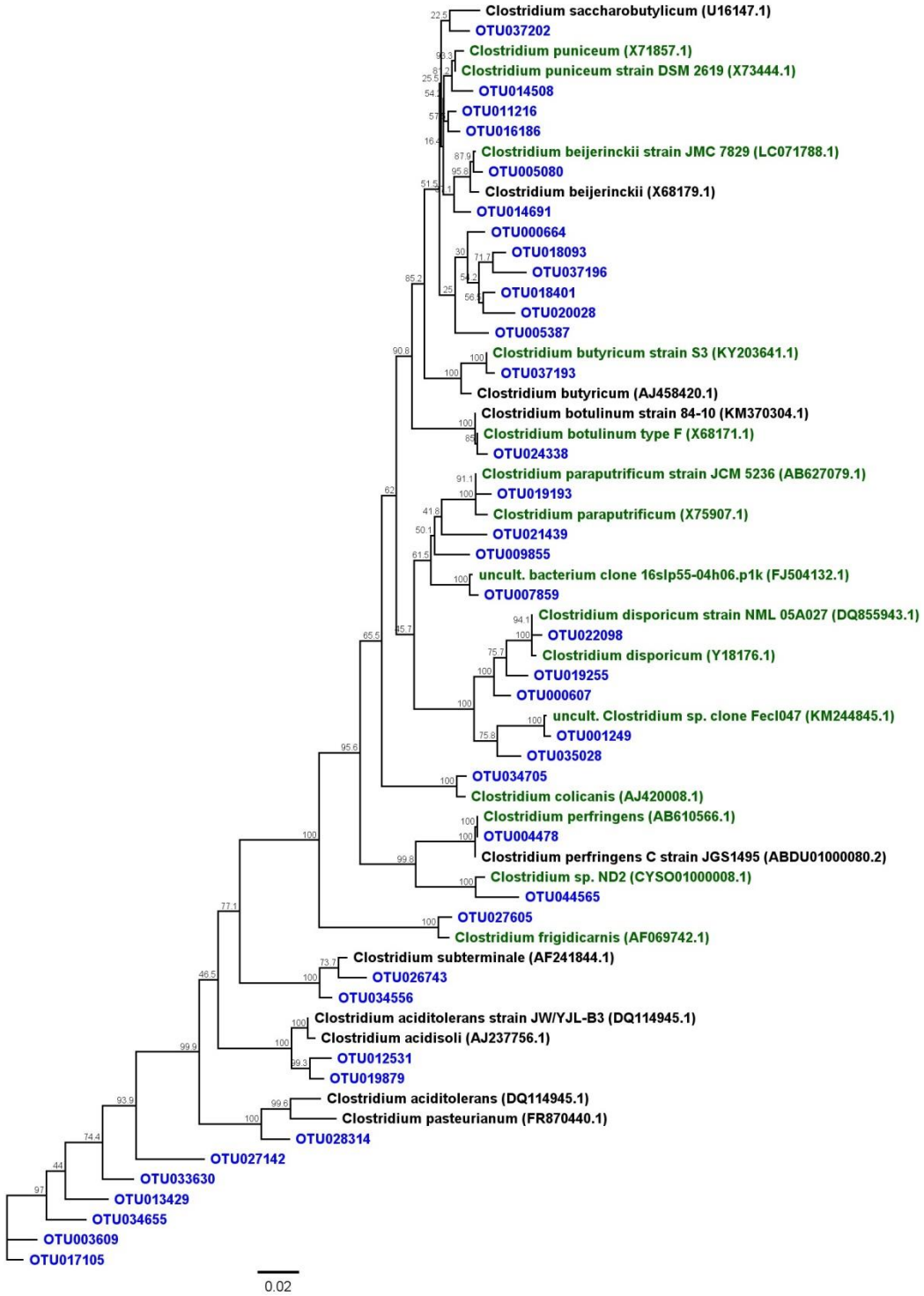
	short-length (477 bp)			**full-lentgh (~ 1,500 bp)**		
	Affiliation	% similarity	Bootstrap	Affiliation	% similarity	Bootstrap
OTU003228	*A. junii*	99.0	39.4	*A. junii*	99.4	100
OTU004317	*A. lwoffii*	98.5	67.3	*A. lwoffii*	98.6	95.3
OTU004358	*A. harbinensis*	98.5	95.1	*A. harbinensis*	98.6	95.9
OTU006694	-	-	-	*A. celticus*	99.7	100
OTU011363	*A. ursingii*	99.0	99.8	*A. ursingii*	99.6	100
OTU012127	*A. albensis*	99.2	99.1	*A. albensis*	99.2	100
OTU012742	*A. baumannii*	98.7	94.6	*A. baumannii*	98.6	100
OTU014283	-	-	-	*A. oleivorans*	98.0	78.0
OTU014486	-	-	-	*A. beijerinckii*	99.7	98.7
OTU017554	*A. schindleri*	98.3	86.3	*A. schindleri*	97.9	98.7
OTU017557	*A. rudis*	96.7	87.9	*A. rudis*	99.3	100
OTU020140	*A. radioresistens*	98.5	94.9	*A. radioresistens*	99.1	100
OTU020183	-	-	-	*A. gandensis*	99.6	100
OTU025639	-	-	-	*A. indicus*	99.1	88.2
OTU025642	*A. gerneri*	96.2	69.5	*A. gerneri*	99.4	100
OTU025938	-	-	-	*A. haemolyticus*	99.1	77.5
OTU036224	*A. beijerinckii*	98.3	49.7	*A. tjernbergiae*	94.0	70.7
OTU037492	*A. haemolyticus*	97.3	66.3	*A. dispersus*	96.9	78.1



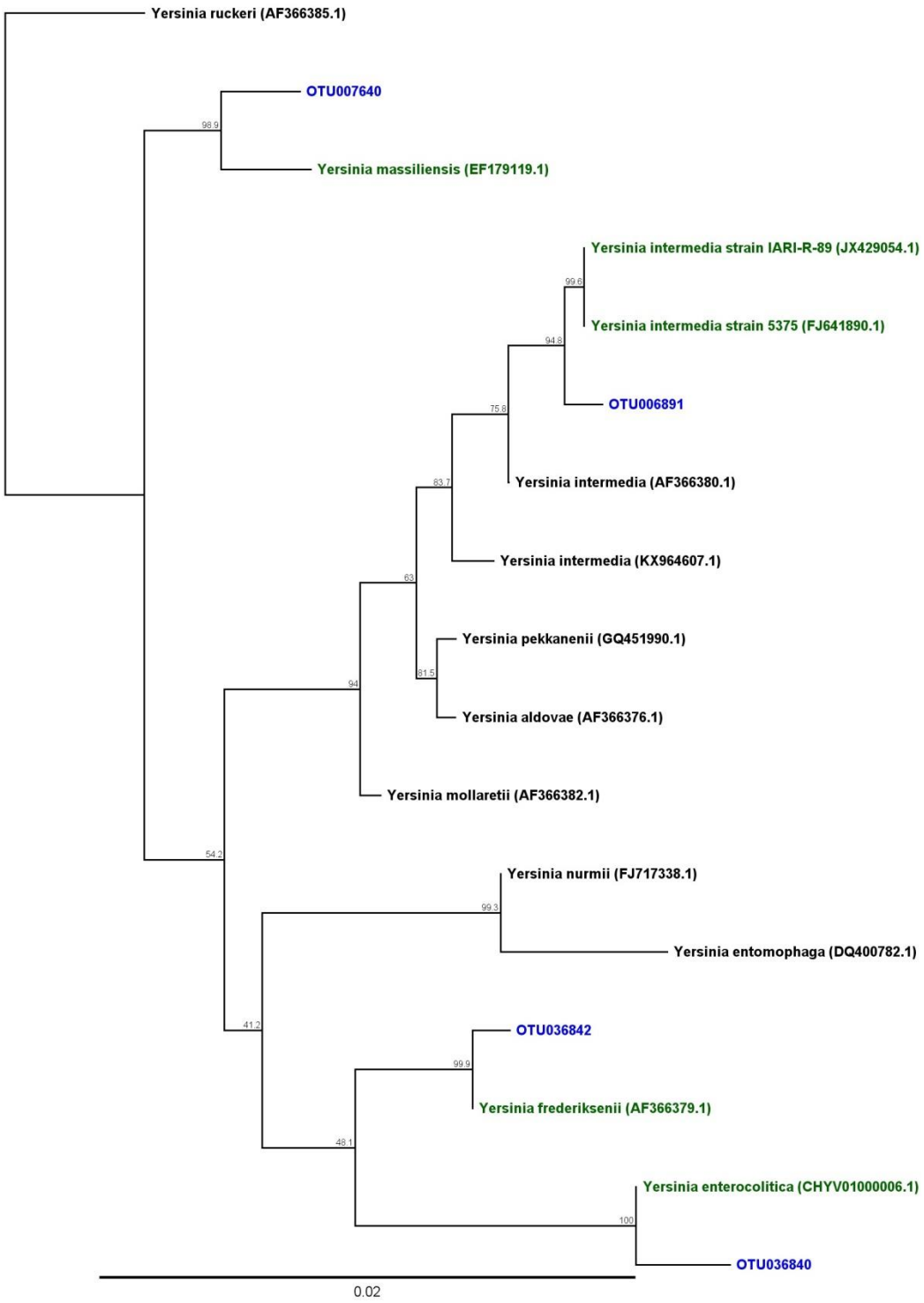
Supplementary Fig. S1: Phylogenetic tree of the genus *Legionella*. Maximum-likelihood tree (PhyML), based on whole 16S rRNA gene or partial sequences over 1,000 bp. Bootstrap percentages at nodes are shown (based on 1,000 iterations). GenBank accession numbers are given in brackets. Bar represents 2% sequence divergence.



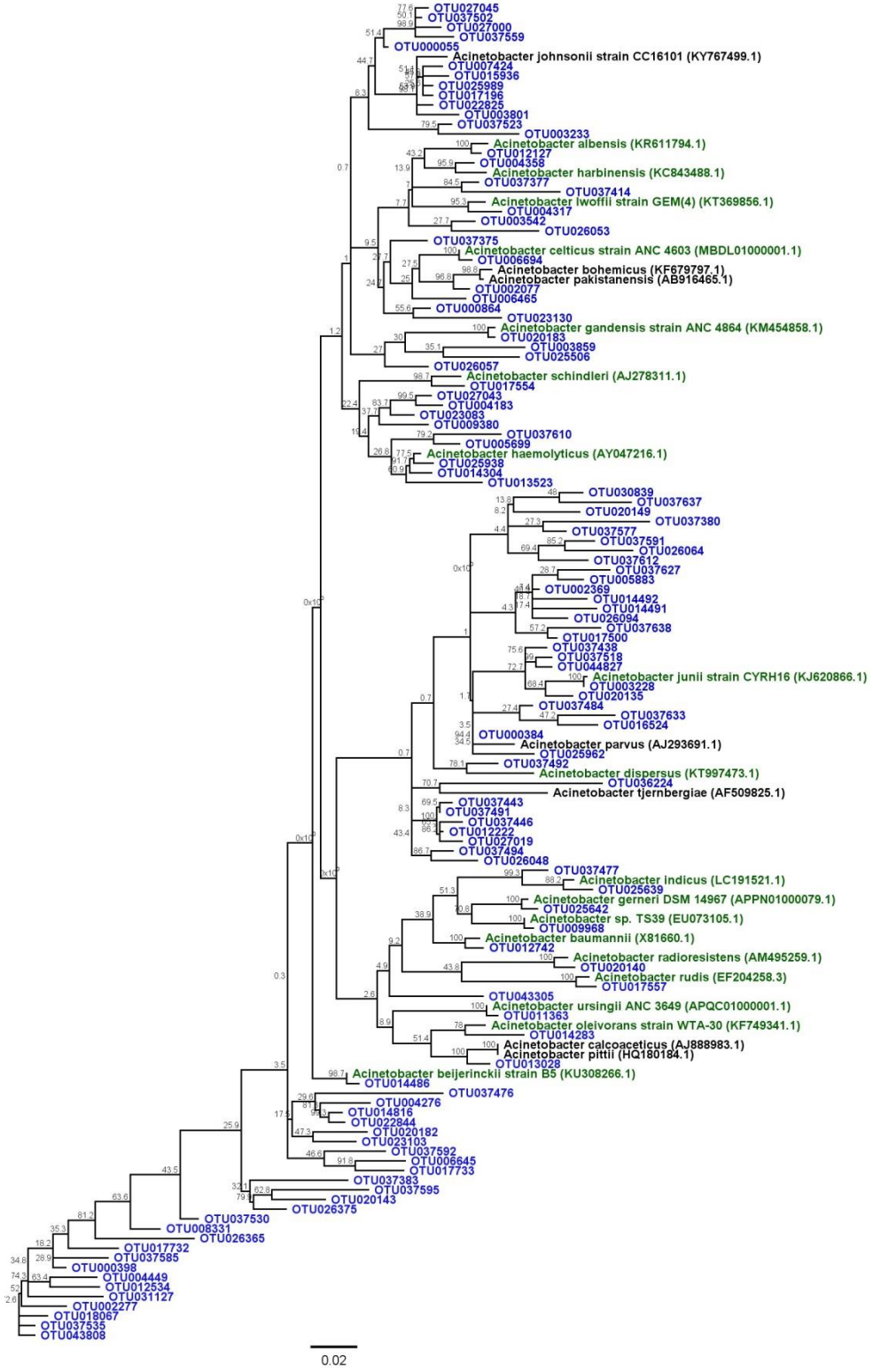
Supplementary Fig. S2: Phylogenetic tree of the genus *Leptospira*. Maximum-likelihood tree (PhyML), based on whole 16S rRNA gene or partial sequences over 1,000 bp. Bootstrap percentages at nodes are shown (based on 1,000 iterations). GenBank accession numbers are given in brackets. Bar represents 2% sequence divergence.



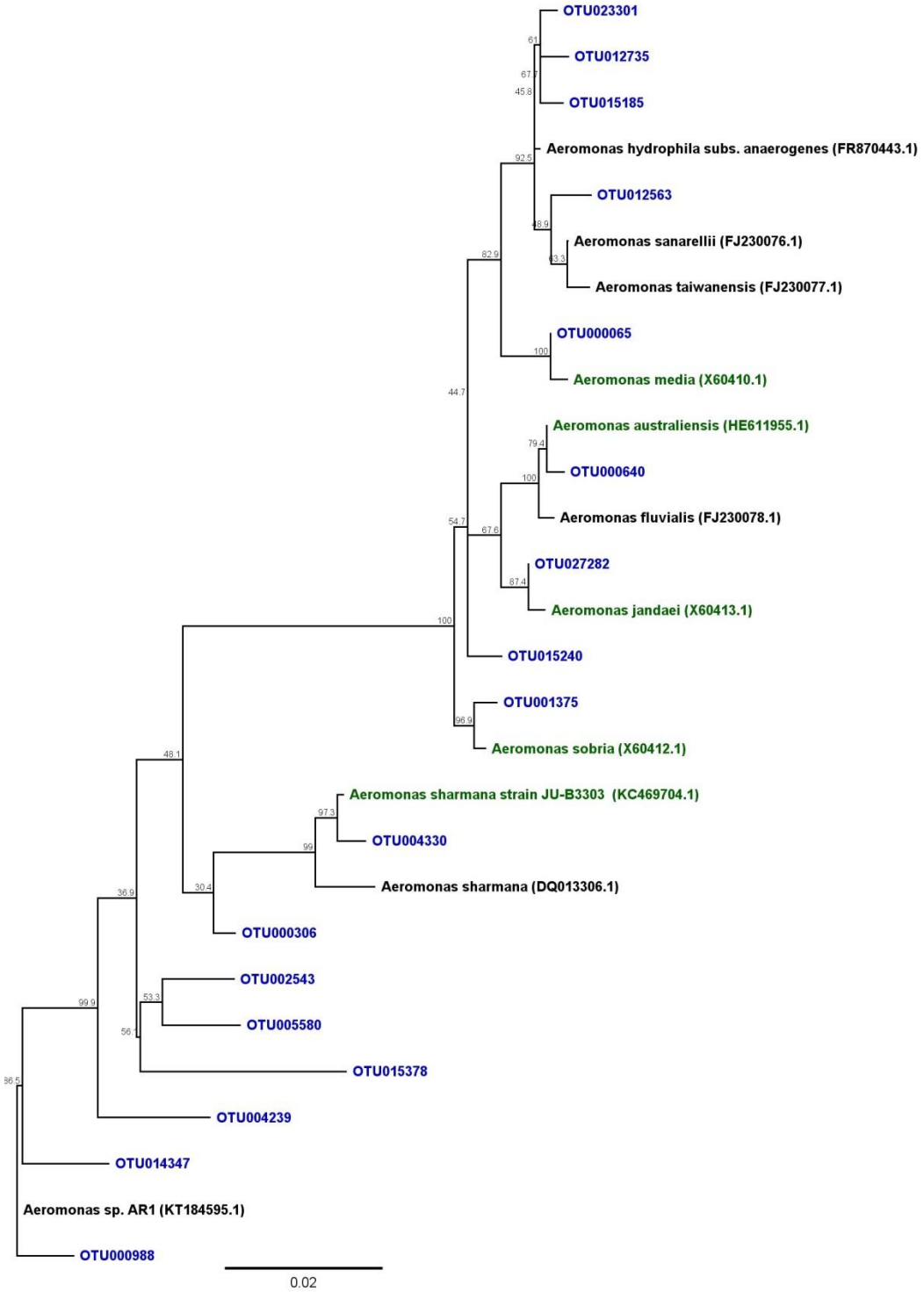
Supplementary Fig. S3: Phylogenetic tree of the genus *Clostridium* (sensu-stricto). Maximum-likelihood tree (PhyML), based on whole 16S rRNA gene or partial sequences over 1,000 bp. Bootstrap percentages at nodes are shown (based on 1,000 iterations). GenBank accession numbers are given in brackets. Bar represents 2% sequence divergence.



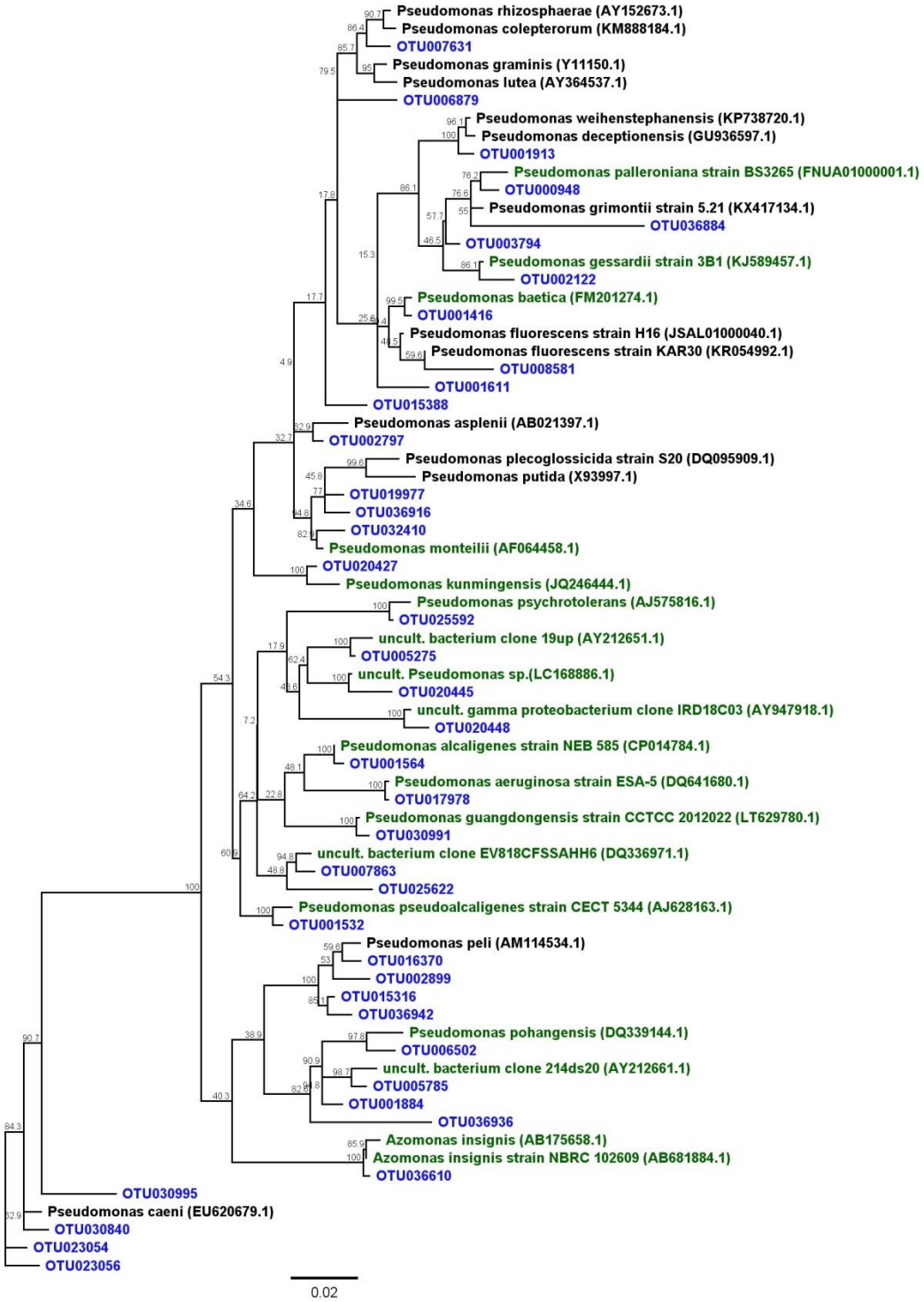
Supplementary Fig. S4: Phylogenetic tree of the genus *Yersinia*. Maximum-likelihood tree (PhyML), based on whole 16S rRNA gene or partial sequences over 1,000 bp. Bootstrap percentages at nodes are shown (based on 1,000 iterations). GenBank accession numbers are given in brackets. Bar represents 2% sequence divergence.



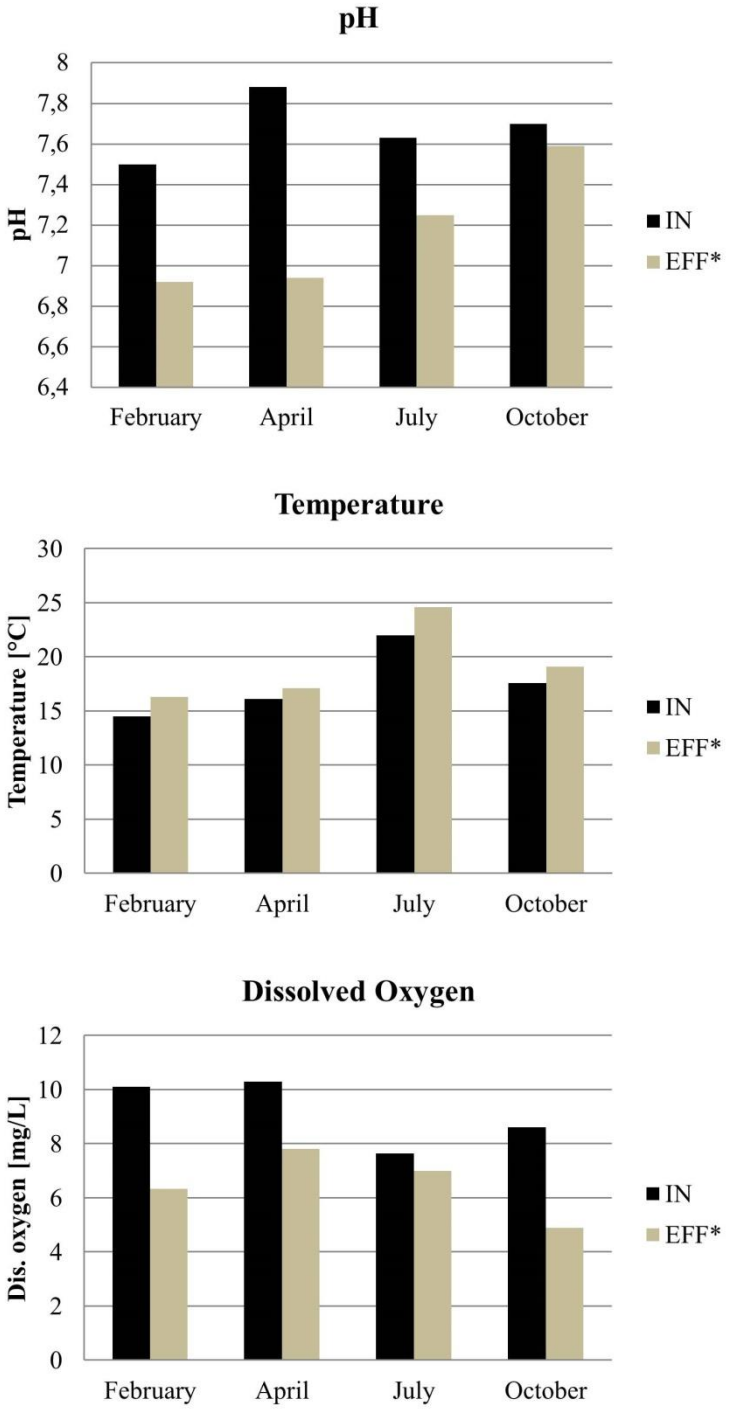
Supplementary Fig. S5: Phylogenetic tree of the genus *Acinetobacter*. Maximum-likelihood tree (PhyML), based on whole 16S rRNA gene or partial sequences over 1,000 bp. Bootstrap percentages at nodes are shown (based on 1,000 iterations). GenBank accession numbers are given in brackets. Bar represents 2% sequence divergence.



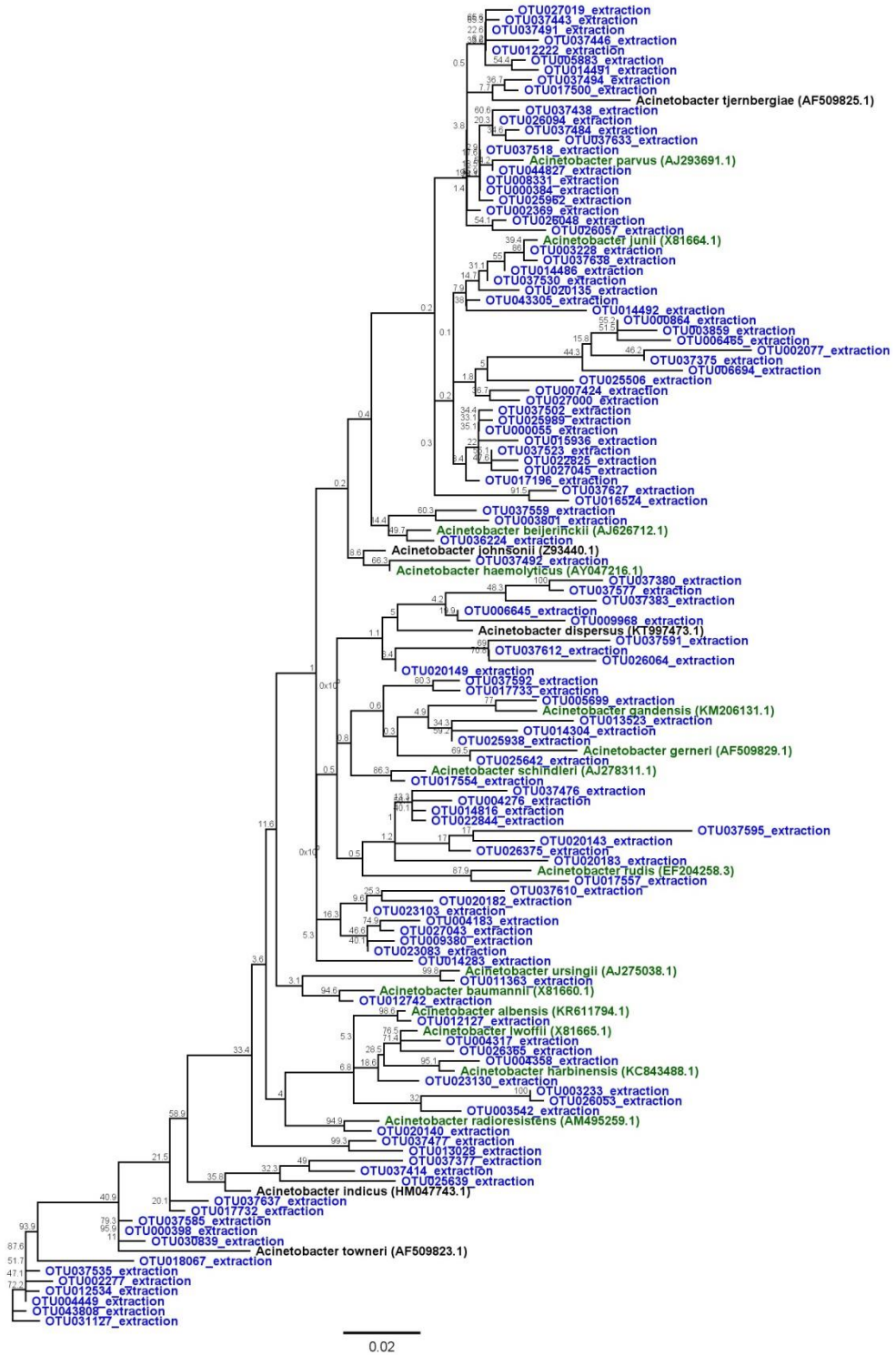
Supplementary Fig. S6: Phylogenetic tree of the genus *Aeromonas*. Maximum-likelihood tree (PhyML), based on whole 16S rRNA gene or partial sequences over 1,000 bp. Bootstrap percentages at nodes are shown (based on 1,000 iterations). GenBank accession numbers are given in brackets. Bar represents 2% sequence divergence.



Supplementary Fig. S7: Phylogenetic tree of the genus *Pseudomonas*. Maximum-likelihood tree (PhyML), based on whole 16S rRNA gene or partial sequences over 1,000 bp. Bootstrap percentages at nodes are shown (based on 1,000 iterations). GenBank accession numbers are given in brackets. Bar represents 2% sequence divergence.



Supplementary Fig. S8: Measurements of pH, temperature and dissolved oxygen from inflow [IN] and effluent [EFF] samples. * Effluent samples were measured by the WWTP staff.



Supplementary Fig. S9: Phylogenetic tree of the genus *Acinetobacter* using 477 bp-sequences extracted from the whole 16S rRNA genes generated in the current study. Maximum-likelihood tree (PhyML), based on whole 16S rRNA gene or partial sequences over 1,000 bp. Bootstrap percentages at nodes are shown (based on 1,000 iterations). GenBank accession numbers are given in brackets. Bar represents 2% sequence divergence.

CHAPTER 2

Genomic analysis of three *Clostridioides difficile* isolates from urban water sources

Genomic analysis of three *Clostridioides difficile* isolates from urban water sources

Daniela Numberger[a], Thomas Riedel[b], Gayle McEwen[a], Ulrich Nübel[b], Martinique Frentrup[b], Isabel Schober[b], Boyke Bunk[b], Cathrin Spröer[b], Jörg Overmann[b], Hans-Peter Grossart[c,d,#], Alex D. Greenwood[a,e,#]

[a] Leibniz Institute for Zoo and Wildlife Research, Alfred-Kowalke-Strasse 17, 10315 Berlin, Germany

[b] Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures, Inhoffenstrasse 7B, 38124 Braunschweig, Germany

[c] Leibniz Institute of Freshwater Ecology and Inland Fisheries, Zur alten Fischerhütte 2, 16775 Stechlin, Germany

[d] University of Potsdam, Institute of Biochemistry and Biology, Karl-Liebknecht-Strasse 24-25, 14476 Potsdam, Germany

[e] Freie Universität Berlin, Department of Veterinary Medicine, Institute for Virology, Robert-von-Ostertag-Strasse 7-13, 14163 Berlin, Germany

[#]**Address correspondence** to Alex D. Greenwood, greenwood@izw-berlin.de and Hans-Peter Grossart, hgrossart@igb-berlin.de

Running head: Genomic analysis of three *C. difficile* isolates

Keywords: *Clostridioides difficile*, genomes, sediment, wastewater

Published in Anaerobe (https://doi.org/10.1016/j.anaerobe.2019.01.002)

ABSTRACT

We investigated inflow of a wastewater treatment plant and sediment of an urban lake for the presence of *Clostridioides difficile* by cultivation and PCR. Among seven colonies we sequenced the complete genomes of three: two non-toxigenic isolates from wastewater and one toxigenic isolate from the urban lake. For all obtained isolates, a close genomic relationship with human-derived isolates was observed.

NOTE

Clostridioides difficile (homotypic synonym *Clostridium difficile*) is a Gram-positive, spore-forming bacterium [1–3]. Pathogenic *C. difficile* strains, mainly defined by their ability to produce the enterotoxin A (TcdA) and/or the cytotoxin B (TcdB) [4–9], can cause gastrointestinal diseases such as diarrhea and life-threatening pseudomembranous colitis, often associated with antibiotic treatment [5, 10–14]. Whereas a role as a nosocomial pathogen is well-documented, evidence is accumulating that animals and non-clinical environments are potential reservoirs for pathogenic *C. difficile* strains and therefore potential sources of infection [10, 15–25]. Genome sequencing data is generally missing from strains isolated from the aquatic environment. Complete genome analyses could clarify the origin, evolution and adaptation of *C. difficile* to the non-clinical environment and provide a better understanding of its epidemiology. Thus, we sequenced the complete genomes of three *C. difficile* strains isolated from the sediment of an urban bathing lake and from the inflow of a wastewater treatment plant (WWTP) in Berlin, Germany using a combination of single molecule real time (SMRT) and Illumina sequencing technology. The obtained genome sequences were further compared to available *C. difficile* genomes.

Inflow samples from a WWTP (Berlin, Germany) were collected on February 11th, April 15th, July 27th and October 20th 2016. A volume of 25-35 mL wastewater was filtered through 0.22 µm Sterivex® filters (Merck Millipore, Germany). The filters were stored at -20°C until DNA extraction according to the manufacturer's instructions using the QIAamp DNA mini kit (Qiagen, Germany). For cultivation, 50 mL of wastewater were centrifuged and

the pellet was stored at -20°C until further processing. The pellet was finally resuspend in 1 mL of sterile Dulbeccos's 1× PBS (Sigma-Aldrich, Germany) for cultivation on chromID® *C. difficile* agar plates (bioMérieux, France) at 37°C under anaerobic condition for at least 48h. Sediment was sampled on August 15th, 2016 from the urban lake 'Weisser See' (52°33'15.3"N, 13°27'48.6"E) in Berlin (Germany), which is heavily used for recreational activity. One gram of sediment was used exclusively for cultivation on chromID® *C. difficile* agar plates.

Colonies were screened for *C. difficile* by a colony PCR on a *C. difficile*-specific region of the 16S rRNA gene. A total of seven colonies originated from the WWTP in spring and summer as well as from the sediment of lake 'Weisser See' were positive and subsequently tested for the presence of the toxin A- and toxin B-encoding gene by PCR. Whereas five obtained isolates from the WWTP inflow were tested negative for toxin A and B, two colonies from the lake were positive. As the isolates from the WWTP were toxin A- and toxin B-negative, we additionally performed a quantitative real-time PCR on the toxin genes (tcdA and tcdB) in combination with a PCR for the *C. difficile*-specific 16S rRNA region using the DNA extracts from the WWTP inflow samples. All PCR conditions, primer and probe sequences used in this study are listed in **Table S1**. To avoid genome sequencing of identical strains or clones we compared PCR-amplified and sequenced fragments of the surface layer protein A gene (slpA) [26, 27]. It revealed three different sequences among our seven colonies, three of which were then selected for complete genome sequencing: DSM 104450 (WWTP spring), DSM 104451 (WWTP summer) and DSM104452 (sediment, 'Weisser See'). For complete genome sequencing, cultivation and genomic DNA extraction was performed as reported previously [28, 29]. Genome sequencing, assembly and annotation was performed as described previously [30]. Briefly, SMRT reads generated on an RSII machine (Pacific Biosciences, USA) were assembled using the RS_HGAP_Assembly.3 protocol implemented in SMRT Portal version 2.3.0 yielding in the complete chromosome and their possibly present extrachromosomal elements of each *C. difficile* isolate. Sequence quality was then improved by mapping Illumina short-reads with >100-fold coverage onto the assembled trimmed and circularized replicons by using BWA (available at https://sourceforge.net/projects/bio-bwa/) [31, 32]. Genomes were annotated by using Prokka 1.8 software (https://github.com/tseemann/prokka) [33]. Sequences were deposited in GenBank under the

accession numbers CP021445.1 + CP021446.1 (extrachromosomal element), CP021447.1 and CP029345.1, respectively.

A search in the *C. difficile* genome sequence database within EnteroBase (http://enterobase.warwick.ac.uk/species/index/clostridium; manuscript in preparation) [34; PMID: 29621240] revealed the following clinical *C. difficile* isolates as close relatives to our environmental isolates: CD-15-00005 (PCR ribotype 010) and CD-15-01025 (PCR ribotype 014) [35], and C00011764 [36]. Data of these strains have been deposited in the European Nucleotide Archive (https://www.ebi.ac.uk/ena) under accession numbers ERR2562463, ERR2562464 and PRJEB4639, respectively. Illumina reads from isolate C00011764 are publicly available (accession number PRJEB4639) [36]. In addition, all 21 available complete genome sequences were downloaded from NCBI GenBank (https://www.ncbi.nlm.nih.gov/genbank/) for comparison (**Table S2**). The amino acid sequences of the protein-coding sequences from the raw Prokka annotations were analysed with Proteinortho v5.16b [37] and genomes were aligned using Parsnp [38]. A phylogenetic tree was constructed using FastTreeDbl and visualised using Figtree [39]. Putative bacteriophages were analysed using the PHASTER web server (http://phaster.ca/, [40]) and BLAST (http//blast.ncbi.nlm.nih.gov/, [41]).

A total of seven *C. difficile* colonies were obtained from WWTP inflow and sediment of an urban lake in Berlin (Germany), of which three were selected for complete genome sequencing. Only the ones derived from the lake were positive for the toxin-encoding genes tcdA and tcdB by PCR. However, quantitative real-time PCR indicated the presence of tcdA and tcdB in WWTP inflow samples, even though the Ct-values were relatively high, which indicate low concentrations probably due to the high dilution in inflow water (**Table S3**). Thus, we assume that both toxigenic (indicated by qRT-PCR) and non-toxigenic strains (isolated) of *C. difficile* were present in WWTP inflow samples throughout the year, likely reflecting fecal contamination [42–45]. Further studies of *C. difficile* in wastewater including its survival in and spread via effluent are required as consistent contamination of the aquatic environment with viable *C. difficile* has already been demonstrated [45, 46].

Phylogenetic analysis revealed that all three isolated *C. difficile* strains belong to clade 1 and show high similarity with human-derived isolates, having only ~5-50 SNPs (single nucleotide

polymorphisms) in regions that could be aligned for all strains. Whereas the WWTP isolates DSM 104450 and DSM 104451 were closely related to isolates CD-15-01025 (PCR ribotype 010) and C00011764 (unknown PCR ribotype), the lake isolate DSM 104452 was closely related to isolate CD-15-00005 (PCR ribotype 014), respectively (**Figure 1**). While CD-15-01025 was collected in Germany, 2014, C00011764 is from United Kingdom. Location and date of collection of strain CD-15-00005 remains unknown. Strains of the same ribotype have already been shown to be isolated along large temporal and geographic distances, as well as from different hosts, which highlights a broad dissemination through many potential transmission routes [47, 48]. General genome features of DSM 104450, DSM 104451, and DSM 104452 are shown in **Table S4**.

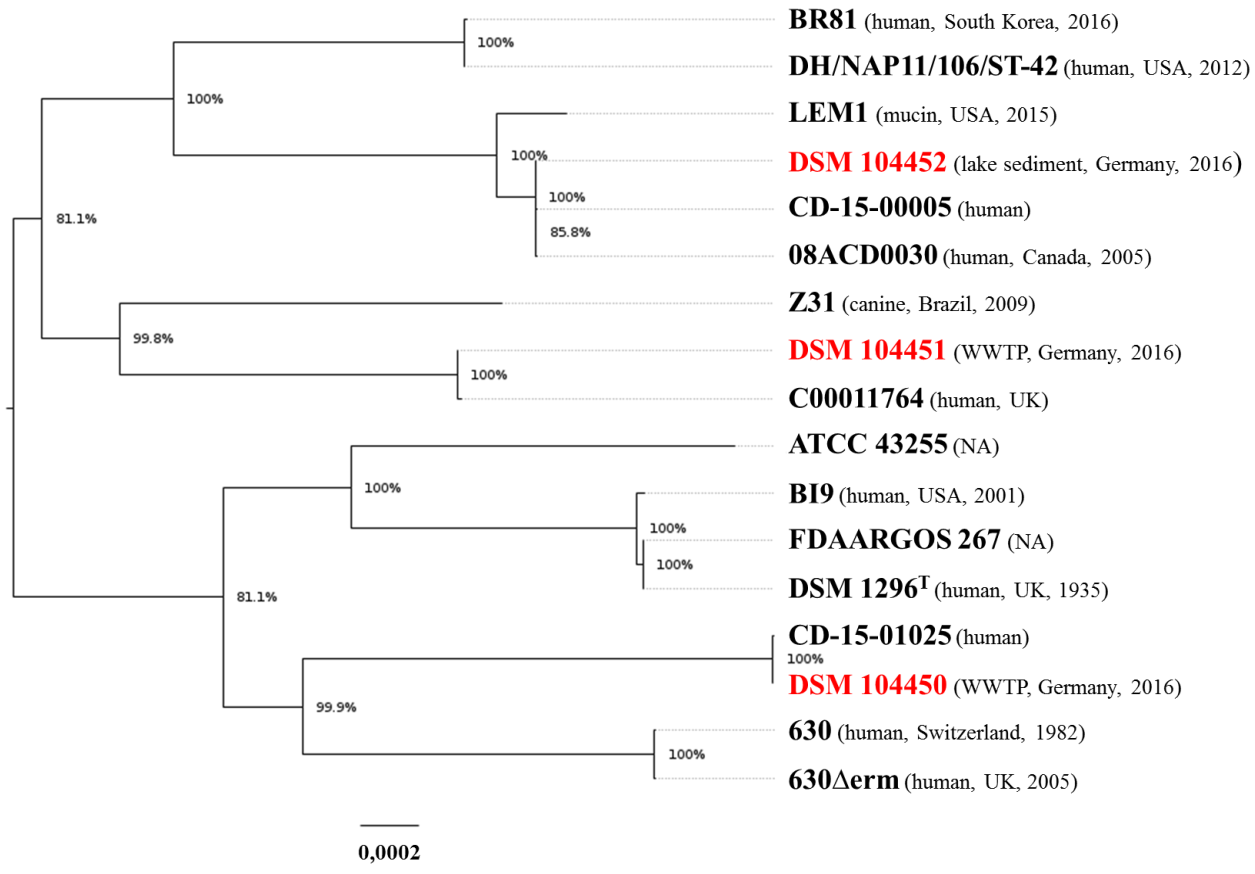


Figure 1: Maximum-likelihood phylogenetic tree of complete sequenced strains of clade 1 based on SNPs (single-nucleotide polymorphisms) using Parsnp [38]. Node labels represent the maximum likelihood bootstrap values calculated from 1,000 repetitions and the scale bar displays 0.003 nucleotide substitutions per site.

The prophage content differed between the three genomes as shown in **Table 1**. Seven prophage sequences, of which four were intact and three incomplete, were identified in DSM 104450. In addition, this isolate contained an extrachromosomal element, which closely matched the Clostridium phages CDSH1, phiCD38-2 and phiCD111. DSM104451 contained two intact and two incomplete prophage sequences. In contrast, DSM104452 had four prophage sequences with only one intact (**Table 1**).

Table 1: Predicted genomic bacteriophage-related regions in three novel *C. difficile* isolates by PHASTER.

Region	Length	Completeness	CDS	Best hit	GC
DSM 104450 (Extrachromosomal element)					
1	41.6 kb	intact	51	*Clostridium* phage phiCD111 [NC_028905]	30.9%
DSM 104450 (Chromosome)					
1	18.1 kb	incomplete	16	*Enterobacteria* phage phi92 [NC_023693]	29.8%
2	68.0 kb	intact	91	Clostridium phage phiMMP03 [NC_028959]	28.5%
3	27.4 kb	incomplete	30	*Clostridium* phage phiCDHM19 [NC_028996]	27.7 %
4	56.0 kb	intact	84	*Clostridium* phage phiMMP01 [NC_028883]	28.6 %
5	45.3 kb	intact	49	*Clostridium* phage phiCD211 [NC_029048]	27.4 %
6	13.3 kb	incomplete	24	*Clostridium* phage phiMMP02 [NC_019421]	25.1%
7	47.4 kb	intact	72	*Clostridium* phage phiCD27 [NC_011398]	29.1%
DSM 104451 (Chromosome)					
1	17.9 kb	incomplete	16	*Enterobacteria* phage phi92 [NC_023693]	29.8%
2	66.3 kb	intact	88	*Clostridium* phage phiMMP03 [NC_028959]	28.7 %
3	27.3 kb	incomplete	30	*Clostridium* phage phiCDHM19 [NC_028996]	28.0%
4	70.7 kb	intact	78	*Clostridium* phage phiMMP03 [NC_028959]	29.5%
DSM 104452 (Chromosome)					
1	18.1 kb	incomplete	16	*Enterobacteria* phage phi92 [NC_023693]	29.8%
2	57.4 kb	intact	90	*Clostridium* phage phiMMP03 [NC_028959]	28.6%
3	27.4 kb	incomplete	30	*Clostridium* phage phiCDHM19 [NC_028996]	27.7%
4	9.4 kb	incomplete	17	*Clostridium* phage phiMMP02 [NC_019421]	26.0%

The pan genome of all 27 analysed complete *C. difficile* genomes consisted of 6,926 CDS and a core genome of 2,620 CDS (38%). A total of 1,513 CDS (22%) were identified as strain unique. DSM 104450, DSM 104451 and DSM104452 had 39, 19 and 2 unique genes, respectively (**Table S5**). *C. difficile* strains with the most unique genes were ATCC 43255, BI9, FDAARGOS_267, LEM1 and Z31. The strains 630, 630Δerm, BI1, BR81, CD196, DH/NAP11/106/ST-42 and DSM 1296T had few (< 10) unique genes. Please note that these numbers are based on only 27 genomes. Thus, future access to more complete genomes will provide more accurate numbers, since upto date most publically available *C. difficile* genome sequences are partial. Previous studies that have included many different strains from different human and animal hosts demonstrated the percentage of conserved functional core genes to be low between 16-19.6% [49] of which 20% encode for hypothetical proteins and the remaining are housekeeping genes for metabolism, biosynthesis, DNA replication, transcription, transport and cell division [49–51]. In the current study, the 38% core gene content is consistent with a number of genes conserved across the investigated strains.

The close relationship of the lake isolate DSM 104452 to the clinical strain CD-15-00005 (PCR ribotype 014) indicates a transmission of this strain between humans and the environment and its success in both of them, wich might particularly occur during intense recreational activity. *C. difficile* PCR ribotype 014 is known to cause infections and is well-established in both human and porcine populations indicating a zoonotic or foodborne etiology [51]. The lake is not connected to any sewer system and is rain fed. Health risk might therefore be highest in summer when recreational activity in and around the lake is most intense.

The low *C. difficile* isolation rate from WWTP inflow could in part be due to the cultivation methods. For example, most samples had to be frozen prior to culturing which may have lowered the isolation rate, though the PCR results suggest that *C. difficile* in general was not extremely abundant in our samples.

Our study provides the first genomic evidence for *C. difficile* in the aquatic environment of Berlin, Germany. To better understand the occurrence and ecology of *C. difficile* in different aquatic environments, further studies will be necessary.

ACKNOWLEDGEMENTS

We thank Vera Junker, Susanne Auls, Elke Mach, Michael Sachtleben, Lars Ganzert, Susanne Schneider, Simone Severitt and Nicole Heyer for technical assistance. We are indebted to Nabil Alikhan, Zhemin Zhou, Martin Sergeant and Mark Achtman (University of Warwick, UK) for establishing the Clostridioides database within EnteroBase. This work was supported by a project grant (IRG 3 - Water) from the Leibniz Research Alliance "INFECTIONS'21 - Transmission Control of Infections in the 21st Century" as funded by the Leibniz Association, Germany (SAS-2015-FZB-LFV). Partial funding was provided by the Federal State of Lower Saxony, Niedersächsisches Vorab (VWZN3215/ZN3266).

COMPETING INTEREST

The authors confirm that they have no conflicts of interest related to the content of this article.

REFERENCES

1. Hafiz S, Oakley CL. *Clostridium difficile*: isolation and characteristics (Plate VIII). J Med Microbiol. 1976; 9:129–36.

2. Hall IC, O'Toole E. Intestinal flora in new-born infants: With a description of a new pathogenic anaerobe, Bacillus difficilis. Am J Dis Child. 1935; 49:390–402.

3. Lawson PA, Citron DM, Tyrrell KL, Finegold SM. Reclassification of *Clostridium difficile* as *Clostridioides difficile* (Hall and O'Toole 1935) Prévot 1938. Anaerobe. 2016;40:95–9.

4. Lyerly DM, Lockwood DE, Richardson SH, Wilkins TD. Biological activities of toxins A and B of *Clostridium difficile*. Infect Immun. 1982; 35:1147–50.

5. Lyerly DM, Saum KE, MacDonald DK, Wilkins TD. Effects of *Clostridium difficile* toxins given intragastrically to animals. Infect Immun. 1985; 47:349–52.

6. Riegler M, Sedivy R, Pothoulakis C, Hamilton G, Zacherl J, Bischof G, et al. *Clostridium difficile* toxin B is more potent than toxin A in damaging human colonic epithelium in vitro. J Clin Invest. 1995; 95:2004–11.

7. Sullivan NM, Pellett S, Wilkins TD. Purification and characterization of toxins A and B of *Clostridium difficile*. Infect Immun. 1982; 35:1032–40.

8. Taylor NS, Thorne GM, Bartlett JG. Comparison of two toxins produced by *Clostridium difficile*. Infect Immun. 1981; 34:1036–43.

9. Warny M, Pepin J, Fang A, Killgore G, Thompson A, Brazier J, et al. Toxin production by an emerging strain of *Clostridium difficile* associated with outbreaks of severe disease in North America and Europe. The Lancet. 2005; 366:1079–84.

10. Båverud V, Gustafsson A, Franklin A, Aspán A, Gunnarsson A. *Clostridium difficile*: prevalence in horses and environment, and antimicrobial susceptibility. Equine Vet J. 2003; 35:465–71.

11. Chen X, Katchar K, Goldsmith JD, Nanthakumar N, Cheknis A, Gerding DN, et al. A mouse model of *Clostridium difficile*-associated disease. Gastroenterology. 2008; 135:1984–92.

12. George RH, Symonds JM, Dimock F, Brown JD, Arabi Y, Shinagawa N, et al. Identification of *Clostridium difficile* as a cause of pseudomembranous colitis. Br Med J. 1978; 1:695–695.

13. Loo VG, Bourgault A-M, Poirier L, Lamothe F, Michaud S, Turgeon N, et al. Host and pathogen factors for *Clostridium difficile* infection and colonization. N Engl J Med. 2011; 365:1693–703.

14. Samie A, Obi CL, Franasiak J, Archbald-Pannone L, Bessong PO, Alcantara-Warren C, et al. PCR detection of *Clostridium difficile* triose phosphate isomerase (*tpi*), toxin A (*tcdA*), toxin B (*tcdB*), binary toxin (*cdtA*, *cdtB*), and *tcdC* genes in Vhembe District, South Africa. Am J Trop Med Hyg. 2008; 78:577–85.

15. Bandelj P, Trilar T, Blagus R, Ocepek M, Rousseau J, Weese JS, et al. Prevalence and molecular characterization of *Clostridium difficile* isolated from European Barn Swallows (Hirundo rustica) during migration. BMC Vet Res. 2014; 10:40.

16. Borriello SP, Honour P, Turner T, Barclay F. Household pets as a potential reservoir for *Clostridium difficile* infection. J Clin Pathol. 1983; 36:84–7.

17. Harvey RB, Norman KN, Andrews K, Hume ME, Scanlan CM, Callaway TR, et al. *Clostridium difficile* in poultry and poultry meat. Foodborne Pathog Dis. 2011; 8:1321–3.

18. Janezic S, Potocnik M, Zidaric V, Rupnik M. Highly Divergent *Clostridium difficile* strains isolated from the environment. PLOS ONE. 2016; 11:e0167101.

19. Kim KH, Fekety R, Batts DH, Brown D, Cudmore M, Silva JJ, et al. Isolation of *Clostridium difficile* from the environment and contacts of patients with antibiotic-associated colitis. J Infect Dis. 1981; 143:42–50.

20. Koene MGJ, Mevius D, Wagenaar JA, Harmanus C, Hensgens MPM, Meetsma AM, et al. *Clostridium difficile* in Dutch animals: their presence, characteristics and similarities with human isolates. Clin Microbiol Infect. 2012; 18:778–84.

21. Pasquale V, Romano VJ, Rupnik M, Dumontet S, Cižnár I, Aliberti F, et al. Isolation and characterization of *Clostridium difficile* from shellfish and marine environments. Folia Microbiol (Praha). 2011; 56:431–7.

22. Pirs T, Ocepek M, Rupnik M. Isolation of *Clostridium difficile* from food animals in Slovenia. J Med Microbiol. 2008; 57 Pt 6:790–2.

23. al Saif N, Brazier JS. The distribution of *Clostridium difficile* in the environment of South Wales. J Med Microbiol. 1996; 45:133–7.

24. Simango C. Prevalence of *Clostridium difficile* in the environment in a rural community in Zimbabwe. Trans R Soc Trop Med Hyg. 2006; 100:1146–50.

25. Simango C, Mwakurudza S. *Clostridium difficile* in broiler chickens sold at market places in Zimbabwe and their antimicrobial susceptibility. Int J Food Microbiol. 2008; 124:268–70.

26. Joost I, Speck K, Herrmann M, Müller L von. Characterisation of *Clostridium difficile* isolates by *slpA* and *tcdC* gene sequencing. Int J Antimicrob Agents. 2009; 33:S13–8.

27. Karjalainen T, Saumier N, Barc M-C, Delmée M, Collignon A. *Clostridium difficile* genotyping based on *slpA* variable region in S-layer gene sequence: an Alternative to Serotyping. J Clin Microbiol. 2002; 40:2452–8.

28. Riedel T, Bunk B, Thürmer A, Spröer C, Brzuszkiewicz E, Abt B, et al. Genome resequencing of the virulent and multidrug-resistant reference strain *Clostridium difficile* 630. Genome Announc. 2015; 3:e00276-15.

29. Riedel T, Bunk B, Wittmann J, Thürmer A, Spröer C, Gronow S, et al. Complete genome sequence of the *Clostridium difficile* type strain DSM 1296T. Genome Announc. 2015; 3:e01186-15.

30. Riedel T, Wetzel D, Hofmann JD, Plorin SPEO, Dannheim H, Berges M, et al. High metabolic versatility of different toxigenic and non-toxigenic *Clostridioides difficile* isolates. Int J Med Microbiol. 2017; 307:311–20.

31. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics. 2009; 25:1754–60.

32. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinforma Oxf Engl. 2010; 26:589–95.

33. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014; 30:2068–9.

34. Alikhan N-F, Zhou Z, Sergeant MJ, Achtman M. A genomic overview of the population structure of *Salmonella*. PLOS Genet. 2018; 14:e1007261.

35. Zaiss NH, Witte W, Nübel U. Fluoroquinolone resistance and *Clostridium difficile*, Germany. Emerg Infect Dis. 2010; 16:675–7.

36. Eyre DW, Cule ML, Wilson DJ, Griffiths D, Vaughan A, O'Connor L, et al. Diverse sources of *C. difficile* infection identified on whole-genome sequencing. N Engl J Med. 2013; 369:1195–205.

37. Lechner M, Findeiß S, Steiner L, Marz M, Stadler PF, Prohaska SJ. Proteinortho: Detection of (Co-)orthologs in large-scale analysis. BMC Bioinformatics. 2011; 12:124.

38. Treangen TJ, Ondov BD, Koren S, Phillippy AM. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. Genome Biol. 2014; 15:524.

39. Rambaut A. FigTree. Tree Figure Drawing Tool. http://tree.bio.ed.ac.uk/software/figtree/. 2009. https://ci.nii.ac.jp/naid/10029324212/. Accessed 16 Apr 2018.

40. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, et al. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res. 2016; 44:W16–21.

41. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990; 215:403–10.

42. Fisher JC, Eren AM, Green HC, Shanks OC, Morrison HG, Vineis JH, et al. Comparison of sewage and animal fecal microbiomes by using oligotyping reveals potential human fecal indicators in multiple taxonomic groups. Appl Environ Microbiol. 2015; 81:7023–33.

43. Kato H, Kita H, Karasawa T, Maegawa T, Koino Y, Takakuwa H, et al. Colonisation and transmission of *Clostridium difficile* in healthy individuals examined by PCR ribotyping and pulsed-field gel electrophoresis. J Med Microbiol. 2001; 50:720–7.

44. Newton RJ, McLellan SL, Dila DK, Vineis JH, Morrison HG, Eren AM, et al. Sewage reflects the microbiomes of human populations. mBio. 2015; 6:e02574-14.

45. Romano V, Pasquale V, Krovacek K, Mauri F, Demarta A, Dumontet S. Toxigenic *Clostridium difficile* PCR ribotypes from wastewater treatment plants in Southern Switzerland. Appl Environ Microbiol. 2012; 78:6643–6.

46. Xu C, Weese JS, Flemming C, Odumeru J, Warriner K. Fate of *Clostridium difficile* during wastewater treatment and incidence in Southern Ontario watersheds. J Appl Microbiol. 2014; 117:891–904.

47. Knetsch CW, Kumar N, Forster SC, Connor TR, Browne HP, Harmanus C, et al. Zoonotic transfer of *Clostridium difficile* harboring antimicrobial resistance between farm animals and humans. J Clin Microbiol. 2018; 56:e01384-17.

48. Stabler RA, Gerding DN, Songer JG, Drudy D, Brazier JS, Trinh HT, et al. Comparative phylogenomics of *Clostridium difficile* reveals clade specificity and microevolution of hypervirulent strains. J Bacteriol. 2006; 188:7297–305.

49. Janvilisri T, Scaria J, Thompson AD, Nicholson A, Limbago BM, Arroyo LG, et al. Microarray identification of *Clostridium difficile* core components and divergent regions associated with host origin. J Bacteriol. 2009; 191:3881–91.

50. Sebaihia M, Wren BW, Mullany P, Fairweather NF, Minton N, Stabler R, et al. The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. Nat Genet. 2006; 38:779–86.

51. Knight DR, Squire MM, Collins DA, Riley TV. Genome analysis of *Clostridium difficile* PCR ribotype 014 lineage in Australian pigs and humans reveals a diverse genetic repertoire and signatures of long-range interspecies transmission. Front Microbiol. 2017; 7. doi:10.3389/fmicb.2016.02138.

52. Frank JA, Reich CI, Sharma S, Weisbaum JS, Wilson BA, Olsen GJ. Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. Appl Environ Microbiol. 2008; 74:2461–70.

53. Page KA, Connon SA, Giovannoni SJ. Representative freshwater bacterioplankton isolated from Crater Lake, Oregon. Appl Environ Microbiol. 2004; 70:6542–50.

54. Lemee L, Dhalluin A, Testelin S, Mattrat M-A, Maillard K, Lemeland J-F, et al. Multiplex PCR targeting *tpi* (triose phosphate isomerase), *tcdA* (toxin A), and *tcdB* (toxin B) genes for toxigenic culture of *Clostridium difficile*. J Clin Microbio. 2004; 42:5710–4.

55. Bélanger SD, Boissinot M, Clairoux N, Picard FJ, Bergeron MG. Rapid detection of *Clostridium difficile* in feces by real-time PCR. J Clin Microbiol. 2003; 41:730–4.

CHAPTER 2: SUPPLEMENTARY MATERIAL

Table S1: Primer and probe sequences plus appropriate PCR reactions and conditions.

Reagent		**Concentration**	**Cycler conditions**		
Whole 16S rRNA gene					
27F	AGR GTT YGA TYM TGG CTC AG			[52, 53]	
1492R	RGY TAC CTT GTT ACG ACTT				
2× MyTaq™ HS Mix (Bioline, UK)		1×	Denaturation	95°C	3 min
20 mg/mL BSA (NEB, USA)		0.56 mg/mL	30 cycles	95°C	30 sec
Primer 27F (10 µM)		0.2 µM		57°C	30 sec
Primer 1492R (10 µM)		0.2 µM		72°C	1 min
Template DNA		variable	Elongation	72°C	3 min
H_2O		variable			
***C. difficile* specific region of 16S rRNA gene**					
Cdiff_16S_F1	TAT TTG AGA GGC ATC TCT T			this study	
Cdiff_16S_R	CCG ATT AAG GAG ATG TCA TTG G				
2× MyTaq™ HS Mix (Bioline, UK)		1×	Denaturation	95°C	3 min
20 mg/mL BSA (NEB, USA)		0.56 mg/mL	30 cycles	95°C	30 sec
Primer Cdiff_16S_F1 (10 µM)		0.2 µM		55°C	30 sec
Primer Cdiff_16S_R (10 µM)		0.2 µM		72°C	1 min
Template DNA		variable	Elongation	72°C	3 min
H_2O		variable			
Colony PCR (toxin genes and *slpA* gene)					
tcdA_F	AGA TTC CTA TAT TTA CAT GAC AAT AT				
tcdA_R	GTA TCA GGC ATA AAG TAA TAT ACT TT			[54]	
tcdB_F	GGA AAA GAG AAT GGT TTT ATT AA				
tcdB_R	ATC TTT AGT TAT AAC TTT GAC ATC TTT				
slpA_com_19-1_F	GTT GGG AGG AAT TTA AGR AAT G			[26]	
slpA_com_20-1_R	GCW GTY TCT ATT CTA TCD TYW				
2× DreamTaq Green PCR Master Mix (Thermo Fisher Scientific, USA)		1×	Denaturation	95°C	10 min
Primer tcdA_F/tcdB_F/slpA_com_19-1_F (10 µM)		0.1 µM	35 cycles	95°C	30 sec
Primer tcdA_R/tcdB_R/slpA_com_20-1_R (10 µM)		0.1 µM		55°C	30 sec
Template DNA		variable		72°C	1 min
H_2O		variable	Elongation	72°C	3 min

Quantitative RT-PCR (toxin genes), Mx3000P cycler (Agilent Technologies, USA)		
tcdA441	TCT ACC ACT GAA GCA TTA C	[55]
tcdA579	TAG GTA CTG TAG GTT TAT TG	
tcdB2667	ATA TCA GAG ACT GAT GAG	
tcdB2746	TAG CAT ATT CAG AGA ATA TTG T	
probe tcdAB1FAM	FAM-5'-CAC GCG GAT TTT GAA TCT CTT CCT CTA GTA GCG CGT G-3'-BHQ1	
probe tcdBB2FAM	FAM-5'-CAC GCC TGG AGA ATC TAT ATT TGT AGA AAC TGG CGT G-3'-BHQ1	

SsoAdvanced™ Universal Probes Supermix (Bio-Rad, USA)	1×	Denaturation	95°C	3 min
20 mg/mL BSA (NEB, USA)	0.7 mg/mL	45 cycles	95°C	10 sec
Primer tcdA441/tcdB2667 (10 µM)	0.5 µM		57°C	30 sec
Primer tcdA579/tcdB2746 (10 µM)	0.5 µM			
Probe tcdAB1FAM/tcdBB2FAM (10 µM)	0.5 µM			
Template DNA	variable			
H_2O	variable			

26. Joost I, Speck K, Herrmann M, Müller L von. Characterisation of *Clostridium difficile* isolates by *slpA* and *tcdC* gene sequencing. *Int J Antimicrob Agents*. 2009; 33:S13–8.

52. Frank JA, Reich CI, Sharma S, Weisbaum JS, Wilson BA, Olsen GJ. Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. *Appl Environ Microbiol*. 2008; 74:2461–70.

53. Page KA, Connon SA, Giovannoni SJ. Representative freshwater bacterioplankton isolated from Crater Lake, Oregon. *Appl Environ Microbiol*. 2004; 70:6542–50.

54. Lemee L, Dhalluin A, Testelin S, Mattrat M-A, Maillard K, Lemeland J-F, et al. Multiplex PCR targeting *tpi* (triose phosphate isomerase), *tcdA* (toxin A), and *tcdB* (toxin B) genes for toxigenic culture of *Clostridium difficile*. *J Clin Microbio*. 2004; 42:5710–4.

55. Bélanger SD, Boissinot M, Clairoux N, Picard FJ, Bergeron MG. Rapid detection of *Clostridium difficile* in feces by real-time PCR. *J Clin Microbiol*. 2003; 41:730–4.

Table S2: *C. difficile* strains selected for phylogenetic and comparative genome analyses.

Clade	Strain	Accession No.	Isolation source/Host	Geographic location and collection year
1	08ACD0030	CP010888.1	*Homo sapiens*	Ontario, Canada, 2005
	CD-15-00005	ERR2562463	*Homo sapiens*	NA
	DSM 104452	CP029345.1	Lake sediment	Berlin, Germany, 2016
	LEM1	CP019469.1	*Mus musculus*	USA, 2015
	BR81	CP019870.1	*Homo sapiens*	South Korea, 2016
	C00011764	ERR347401	*Homo sapiens*	United Kingdom
	DSM 104451	CP021447.1	WWTP	Berlin, Germany, 2016
	DSM 27543 (630)	CP010905.2	*Homo sapiens*	Switzerland, 1982
	DSM 28645 (630Δ*erm*)	CP016318.1	*Homo sapiens*	United Kingdom, 2005
	Z31	CP013196.1	*Canis familiaris*	Brazil, 2009
	BI9	FN668944.1	*Homo sapiens*	USA, 2001
	DSM 1296^{T}	CP011968.1	*Homo sapiens*	England, 1935
	ATCC43255	CM000604.1	NA	NA
	FDAARGOS_267	CP020424.1	NA	NA
	DH/NAP11/106/ST-42	CP022524.1	*Homo sapiens*	Chicago, USA, 2012
	CD-15-01025	ERR2562464	*Homo sapiens*	Germany, 2014
	DSM 104450	CP021445.1	WWTP	Berlin, Germany, 2016
2	BI1	FN668941.1	*Homo sapiens*	USA, 1988
	CD196	FN538970.1	*Homo sapiens*	NA
	CIP 107932	CM000659.1	*Homo sapiens*	Marne, Reims, France
	2007855	FN665654.1	*Bos taurus*	USA, 2007
	R20291	FN545816.1	*Homo sapiens*	United Kingdom
4	BJ08	CP003939.1	*Homo sapiens*	Beijing, China, 2008
	M68	FN668375.1	*Homo sapiens*	Ireland, 2006
	CF5	FN665652.1	*Homo sapiens*	Belgium, 1995
5	M120	FN665653.1	*Homo sapiens*	United Kingdom, 2007
	QCD-23m63	CM000660.1	*Homo sapiens*	Quebec, Montreal, Canada

Table S3: Ct-values of quantitative real-time PCR on toxin A (*tcdA*) and toxin B (*tcdB*) genes (high Ct-values correspond to low concentration) and result of conservative PCR on the *C. difficile*-specific 16S rRNA gene sequence (✓ indicates a positive result by having a band and subsequently confirmed as *C. difficile* by sanger sequening).

	Ct-Values						Bands
	tcdA			***tcdB***			***16S rRNA***
Triplicate	1	2	3	1	2	3	-
WWTP (11/2/2016)	33.9	no Ct	36.7	37.8	38.0	37.4	✓
WWTP (15/4/2016)	no Ct	no Ct	no Ct	no Ct	no Ct	no Ct	✓
WWTP (27/7/2016)	39.0	37.7	38.5	no Ct	38.8	no Ct	✓
WWTP (20/10/2016)	34.1	37.8	34.9	37.4	38.4	36.9	✓

Table S4: Genome statistics of three novel *C. difficile* isolates from Berlin.

Isolate	**Post-filtered reads**	**Coverage**	**Genome size (bp)**	**GC content**	**No. of coding sequences**	**No. of tRNA genes**	**No. of rRNA genes**
DSM 104450	72,836	123x	4,229,131	28.9%	3,802	90	35
			41,619*	30.9%	51	0	0
DSM 104451	78,573	137x	4,164,224	29.0%	3,697	90	35
DSM 104452	76,054	150x	4,167,361	28.8%	3,661	90	35

* extrachromosomal element

Table S5: Unique coding sequences (CDS) of three novel *C. difficile* isolates from Berlin compared to the complete genome sequences of 24 *C. difficile* strains.

Locus Tag	Protein
DSM 104450 (WWTP)	
CDIF104450_00551	DNA-binding protein
CDIF104450_00733	Putative signal peptide
CDIF104450_02138	Uridine phosphorylase
CDIF104450_02139	Putative transcriptional regulator
CDIF104450_04035	Single-strand binding protein
CDIF104450_04040	Resolvase
CDIF104450_04041	Phage terminase small subunit
CDIF104450_04042	Phage terminase, large subunit, PBSX family
CDIF104450_04043	Bacteriophage portal protein, SPP1 Gp6-like protein
CDIF104450_04044	Minor capsid 2 protein
CDIF104450_04047	Minor structural GP20 protein
CDIF104450_04048	Phage coat protein
CDIF104450_04051	Minor capsid
CDIF104450_04055	Protein gp15
CDIF104450_04056	Tetratricopeptide repeat protein
CDIF104450_04057	Phage tail tape measure protein, TP901 family
CDIF104450_04059	Gp14 protein
CDIF104450_04061	Phage pre-neck appendage-like protein
No. of hypothetical proteins	21
DSM 104451 (WWTP)	
CDIF104451_00713	Putative signal peptide
CDIF104451_03080	DNA-binding protein
CDIF104451_03105	Phage anti-repressor
CDIF104451_03353	Conjugative transposon protein
CDIF104451_03359	Two-component system response regulator
CDIF104451_03362	Putative signal peptide
CDIF104451_03369	Integrase
CDIF104451_03370	Transporter
No. of hypothetical proteins	11
DSM 104452 (Lake)	
CDIF104452_00682	Putative signal peptide
CDIF104452_03333	Ribosome biogenesis GTPase YqeH
No. of hypothetical proteins	0

CHAPTER 3

Recovery of influenza A viruses from lake water and sediments by experimental inoculation

Recovery of influenza A viruses from lake water and sediments by experimental inoculation

Daniela Numberger[1], Carola Dreier[2,3], Colin Vullioud[1], Gülsah Gabriel[2,4], Alex D. Greenwood[1,6]*, Hans-Peter Grossart[5,7]*

1 Leibniz Institute for Zoo and Wildlife Research, Alfred-Kowalke-Straße 17, 10315 Berlin, Germany

2 Heinrich Pette Institute, Leibniz Institute for Experimental Virology, Martinistraße 52, 20251 Hamburg, Germany

3 current address: University of Ulm, Institute of Neurobiology, Albert-Einstein-Allee 11, 89081 Ulm, Germany

4 University of Veterinary Medicine Hannover, Foundation, Bünteweg 2, 30559 Hannover, Germany

5 Leibniz Institute of Freshwater Ecology and Inland Fisheries, Zur alten Fischerhütte 2, 16775 Stechlin, Germany

6 Freie Universität Berlin, Department of Veterinary Medicine, Institute for Virology, Robert von Ostertag-Straße 7-13, 14163 Berlin, Germany

7 University of Potsdam, Institute of Biochemistry and Biology, Karl-Liebknecht-Straße 24-25, 14476 Potsdam, Germany

***Correspondence: Hans-Peter Grossart** (hgrossart@igb-berlin.de) and **Alex D. Greenwood** (greenwood@izw-berlin.de)

Running title: Recovery of influenza A viruses

Submitted and under revision (in PLOS ONE).

ABSTRACT

Influenza A viruses (IAV) are zoonotic pathogens relevant to human, domestic animal and wildlife health. Many avian IAVs are transmitted among waterfowl via a faecal-oral-route. Therefore, environmental water where waterfowl congregate may play an important role in the ecology and epidemiology of avian IAV. Water and sediment may sustain and transmit among individuals or species.

It is unclear at what concentrations waterborne viruses are infectious or remain detectable. To address this, we performed lake water and sediment dilution experiments with varying concentrations of four IAV strains from seal, turkey, duck and gull. To test for infectivity of the IAV strains in a concentration dependent manner, we applied cultivation to specific pathogen free (SPF) embryonated chicken eggs and Madin-Darby Canine Kidney (MDCK) cells.

IAV recovery was more effective in embryonated chicken eggs than MDCK cells for freshwater lake dilutions, whereas, MDCK cells were more effective for viral recovery from sediment samples. Low virus concentration (1PFU/200 µL infection volume) was sufficient in most cases to detect and recover IAV from lake water dilutions, whereas in sediment higher viral concentrations are required to seed an infection.

Keywords: Influenza A, environmental samples, cultivation, detection limit, water, sediment

INTRODUCTION

Influenza A viruses (IAV) are widespread single stranded negative-sense RNA viruses with a broad host range including birds (1–3), humans (4–7), horses (8–10), pigs (11–13) and marine mammals (14,15). Waterfowl are the natural reservoirs of IAVs (2,16,17) and avian IAV can cause fatal outbreaks among wild birds and poultry (16–22). Human infections with avian IAVs demonstrate the zoonotic potential of IAVs (20,23–31).

Avian IAVs are shed into water by birds in high concentration via faeces (3,32,33). Once shed, IAVs remain both environmentally persistent and infectious, particularly in cold freshwater (4°C, 17°C) with 0 ppt salinity (3,34,35). The role of the environment in IAV transmission, particularly water sources used regularly by waterfowl and other bird species is not fully understood. However, there is increasing evidence that water plays an important role in the ecology, epidemiology and transmission of avian IAV (3,24,32,36,37). Whereas human and other mammalian IAVs are mainly transmitted through smear infection and inhalation of aerosols and droplets (4,38–40), avian IAVs are transmitted via a faecal-oral-route. Water may play an important role in indirect transmission via faecal contamination (3,33,37,41).

Leung et al. (42) indicated that water such as from poultry drinking troughs, can be used in avian influenza surveillance . Water can be both simultaneously contaminated by multiple strains and infect multiple individuals. This is also true for natural water bodies which are often heavily used by waterfowl and consequently contaminated by IAVs. This has been shown for a lake used by an Alaskan dabbling duck population (32) and water bodies along the Atlantic Flyway (43). Thus, it could be quite useful to include lake water and sediment when conducting IAV surveillance of wild waterfowl populations and drinking water when working with any kind of poultry.

However, the detection of infectious IAVs in water is challenging. There are no standardized methods for the detection and isolation of IAVs from water and sediment samples. In addition to the expected high dilution of virus in water bodies detection from sediments is complicated by high concentrations of microbes and substances that can interfere with viral culturing experiments, e.g. bacteria and fungi. Sediment might contain higher viral concentrations due to sedimentation processes and the virus might be protected there (e.g.

from UV-light) resulting in longer persistence (44–46). This has been demonstrated in an experiment in which lake sediment, duck feces and duck meat was inoculated with IAV and persistence was found to be longest in lake sediment (46).

Sixty seven sediment samples from five lakes were screened by influenza virus specific PCR with 26 (= 38.8%) found to be positive (Supplementary Table S1). Viral isolation attempted from the PCR positive samples using embryonated chicken eggs and Madin-Darby Canine Kidney (MDCK) cell cultures was, however, unsuccessful. To determine if the lack of cultivation success was due to low viral concentration, dilution experiments were performed with water and sediment using four distinct serially diluted IAV strains with addressing the hypothesis that there are differences in the probability of recovering IAV from freshwater and sediment samples using different predictors (initial virus concentration, method and virus strain). Embryonated eggs and MDCK cell cultures were then performed on the diluted strains to determine the minimal viral concentration needed for IAV detection by cultivation and the results were statistically evaluated. The results are discussed in the context of the persistence of IAVs in the environment and the applicability of water and sediment to IAV surveillance.

MATERIALS AND METHODS

The experimental design is shown schematically in **Figure 1**. The experiment was designed according to the protocols of virus isolation from beach sand and sediment by Poulson et al. (47) and Dalton et al. (43), respectively. We additionally compared our experimental design to a similar study in which they established a protocol to concentrate and recover influenza A viruses from large volumes of water (48). Instead of using a non-pathogenic reverse-genetic virus we decided to use four different pathogenic virus strains. We used five dilutions and applied embryonated chicken egg and MDCK cell culturing as they did to obtain sufficient data.

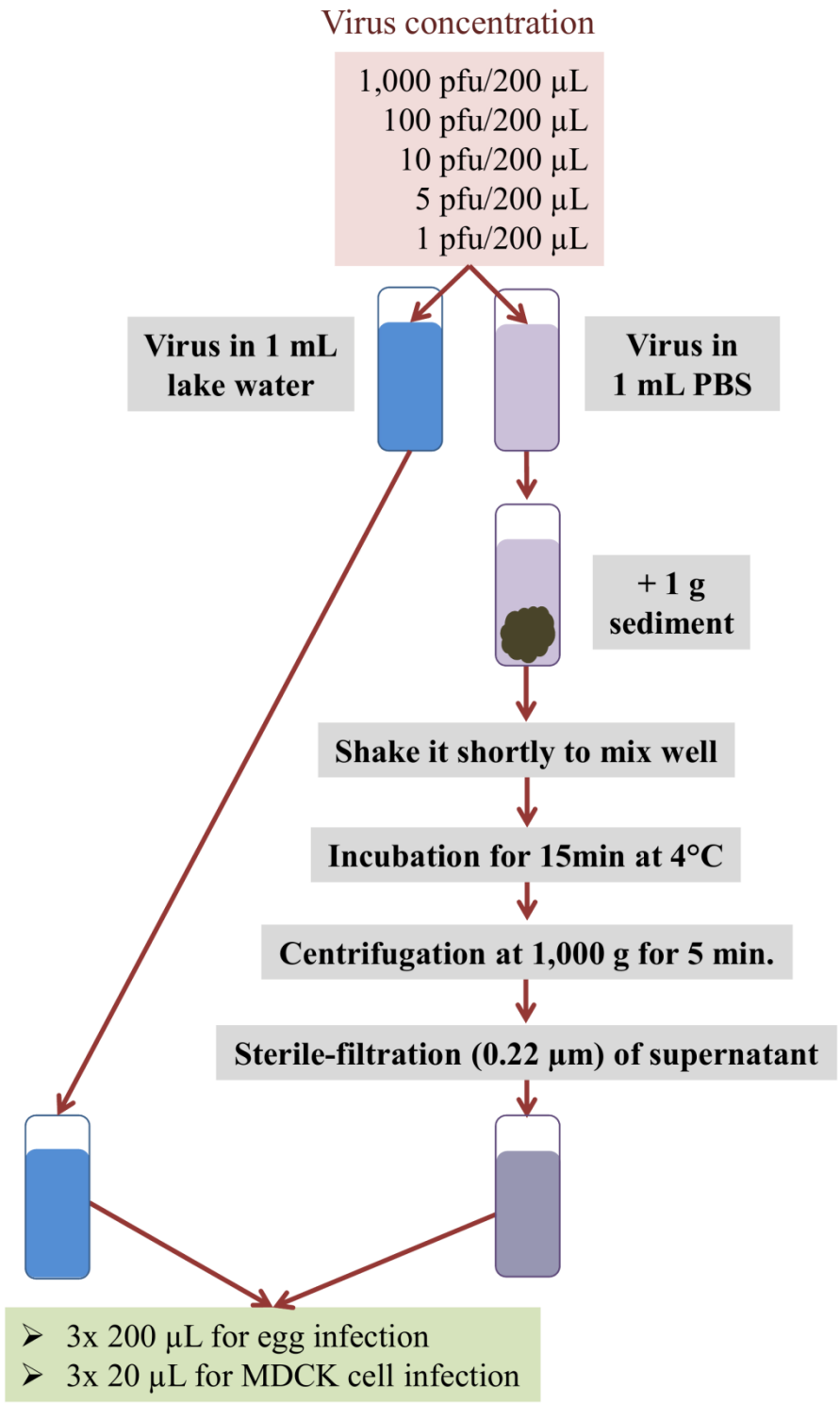


Figure 1: Schematic representation of the dilution experiment design. Lake surface water and sediment were collected on October 16, 2015 from the lake Stechlin (Brandenburg, Germany) and influenza A virus dilutions were made with the collected water. Detection and recovery from inoculated water and sediment samples were performed by cultivation in embryonated chicken eggs and Madin-Darby Canine Kidney (MDCK) cells.

Ethics Statement

According to the German animal protection law experiments with (11-days) embryonated chicken eggs do not need any specific permission. No animal experiments were performed.

Sampling of lake water and sediment

Water and sediment samples used for the inoculation experiment were taken from oligo-mesotrophic lake Stechlin (53°08'56.7", N 13°02'33.5"E) on October 16, 2015. Surface water was collected with sterile 50 mL tubes and the first centimetre of the sandy sediment was obtained using a plexiglas tube (length 50 cm, Ø 44 mm) as a sediment corer. Samples were stored at -20°C. Both sample types were tested for IAV by conventional PCR of a 104 bp size fragment of the Matrix gene according to Ward et al. (49). After reamplification and a total of 70 amplification cycles (Supplementary **Table S1**) all samples were negative. The lake water was sterile-filtered through 0.22 µm Sterivex® filters (Merck Millipore, Germany) before use. The sediment sample was mixed with a sterile spoon and separated in twenty 1 g aliquots (for each strain and dilution).

Influenza A strains and dilutions

The experiment was set up in such a way that IAV negative natural water and sediment representing realistic conditions in which IAVs are transmitted, e.g. an environment in which wild and domestic bird influenza outbreaks occur were used. Experiments were carried out under S3 laboratory condition on the Heinrich Pette Institute. The following IAV strains were tested: A/Seal/Massachusetts/1/1980 (H7N7) (50), A/Turkey/England/1977 (H7N7) (51), A/Gull/Maryland/704/1977 (H13N6) (52) and A/Duck/Alberta/35/1976 (H1N1) (53). The viral stocks were diluted in 1 mL sterile filtered lake water to obtain 1000, 100, 10, 5 and 1 plaque forming units (PFU) per 200 µL (volume used for inoculation). Virus dilutions that were used to spike sediment samples were prepared each in a final volume of 1 mL as described above but in sterile Dulbeccos's 1× PBS (Sigma-Aldrich, Merck KGaA, Germany) and then added to the 1 g sediment aliquots and mixed thoroughly by shaking and inverting the tube. Samples that were not inoculated with influenza A viruses served as negative controls and showed no cytopathogenic effects in MDCK cells nor hemagglutination with 1%

chicken erythrocyte suspension. We did not include additional positive controls as we added pure virus the samples and the strains were tested before when determining their stock concentrations.

IAV cultivation in embryonated chicken eggs

Specific pathogen free (SPF) eggs were obtained from VALO BioMedia GmbH (Germany) and incubated at 37°C and 55-60% rH for 11 days. Each infection was performed in triplicates (three eggs per sample, for each virus strain and virus concentration) under BSL-3 conditions at the Heinrich Pette Institute, Leibniz Institute for Experimental Virology in Hamburg. 200 µL of each virus dilution was used for infection of eggs. The infected amino-allantoic fluid was then incubated at 37°C for 48 h and analysed subsequently for viral replication by hemagglutination assay.

Hemagglutination assays

The hemagglutination assay was performed according to Hirst (54) with some modifications. For the assay, 1% chicken erythrocyte suspension was prepared in 0.9% sodium chloride (Th. Geyer GmbH & Co.KG, Germany). Chicken blood was purchased from Lohmann Tierzucht, Cuxhaven, Germany. Spiked samples, negative and positive controls were diluted 2-fold with Dulbeccos's 1× PBS (Sigma-Aldrich, Merck KGaA, Germany). 50 µL of 1% erythrocyte solution was added to each virus dilution in a 96-well V-bottom microtiter plate. Hemagglutination was evaluated after incubation for 30 min at 4°C by checking each well for agglutination of red blood cells.

MDCK cell culture and cytopathogenic effect measurements

A continuous line of Madin Darby canine kidney II (MDCK II) cells was grown in minimal essential medium (MEM, Gibco, Gibco Life Technologies, Germany) supplemented with 10% fetal bovine serum (Invitrogen, Thermo Fisher Scientific, USA), 1% L-Glutamin (Sigma-Aldrich, Merck KGaA, Germany), 1% Penicillin und Streptomycin (Sigma-Aldrich, Merck KGaA, Germany). Cells were infected as described before (modified after Gaush and Smith (55)). Infection of MDCK cells was performed at 37°C for 48 hours in

96-well microtiter plates containing infection medium (MEM with 1% L-Glutamin, 0.2% BSA, 1% each Penicillin, Streptomycin and 1 mg/mL TPCK-trypsin (Sigma-Aldrich, Merck KGaA, Germany). The viruses were serially diluted to obtain 1, 5, 10, 100 and 1000 PFU. Cytopathic effects were evaluated by light microscopy.

Statistical Analysis

The analyses were performed in R version 3.5.2 (56). We fitted a GLM logistic regression to predict the probability of recovery of IAV and to assess the effect of the sample type, the method, the strains and the initial concentration of virus. We added an interaction between the sample type and the method to account for the differential effect of the used method in different environmental conditions to the model. The model was fitted with the function 'fitme' of the spaMM package (57,58) using penalized quasi-likelihood (using Method = "PQL", in the function call). P-values were obtained through likelihood ratio test between the full model and models, in which the effect under investigation was removed (using the function 'spaMM::anova()').

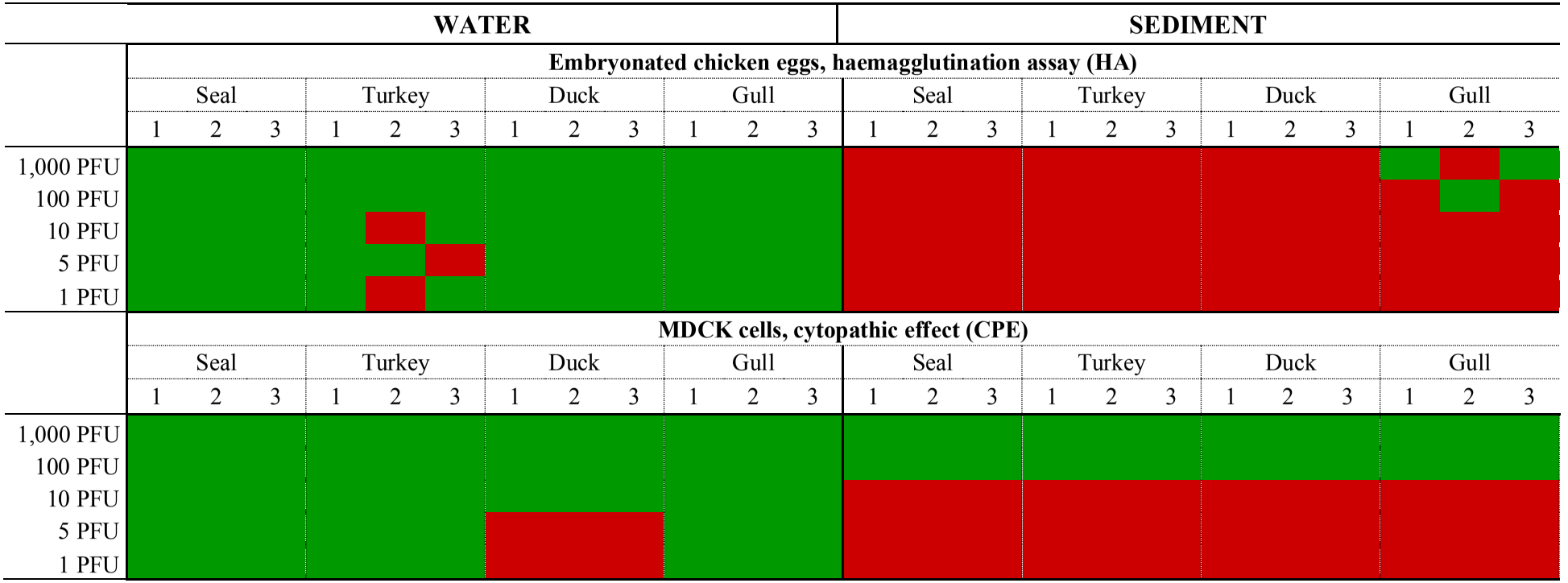
RESULTS

To ensure that the hemagglutination and cytopathogenic effects observed using experimental IAV dilutions were caused by the introduced IAV laboratory strains and not by viruses in the samples themselves, non-inoculated lake water and sediment samples were included as negative controls. The negative controls did not show any hemagglutination or cytopathogenic effects. They were also all PCR negative. The cultivation results for virus recovery from diluted IAV freshwater and sediment samples are summarized in **Table 1**.

Sample type

Generally, virus recovery by culture from sediment was significantly less efficient than from freshwater ($\chi 2 = 174.69$, $p < 0.001$, **Figure 2A**, **Figure 3**) using both embryonated chicken eggs (log-odd = 11.18) and MDCK cell cultures (log-odd = 3.17).

Table 1: Results of spiking experiment to test recovery rate of different IAV strains from water and sediment samples. The cultivation in embryonated chicken eggs was evaluated by the hemagglutination assay. Infection in MDCK cells was shown as cytopathic effect. Green and red colours indicate positive and negative results, respectively.



Cultivation method

MDCK cells and embryonated chicken eggs have different effects on the probability of recovery in freshwater and in sediment ($\chi 2 = 27.68934$, $p < 0.001$, **Figure 2A**, **Figure 3**). In freshwater, embryonated chicken eggs were more efficiently for recovery than MDCK cells (log-odd = 1.23), while for sediment samples the MDCK cells worked better (log-odd = 6.78).

For sediment samples, all infected eggs, except for three eggs infected with the A/Gull/Maryland/704/1977 (H13N6) strain, were negative, whereas recovery was successful in MDCK cell cultures, when having ≥ 100 PFU/100 µL (**Table 1**, **Figure 2A**).

For freshwater samples, both cultivation systems worked good with a higher probability of recovery in embryonated chicken eggs, as the duck strain A/Duck/Alberta/35/1976 (H1N1) was not recovered from starting concentrations lower than 10 PFU/200 µL when using MDCK cells (**Table 1**, **Figure 2**).

Strains and initial virus concentration

Strains ($\chi 2 = 12.077$, $p = 0.007$) and initial virus concentration ($\chi 2 = 39.512$, $p > 0.001$) also had significant effects on the probability of recovering the virus (**Figure 3**). Recovery from sediment samples was only successful when having starting viral concentrations of ≥ 100 PFU/200 µL, whereas in freshwater samples also 1 PFU/200 µL could be sufficient to be recovered. In general, the probability of recovery increased with higher starting concentrations.

A different probability of recovery was observed for the strains A/Gull/Maryland/704/1977 (H13N6) and A/Duck/Alberta/35/76 (H1N1) as A/Duck/Alberta/35/1976 (H1N1) showed a more efficient recovery in embryonated chicken eggs than in MDCK cells and A/Gull/Maryland/704/1977 (H13N6) was the only strain that had a small chance to be recovered from sediment samples when using embryonated chicken eggs.

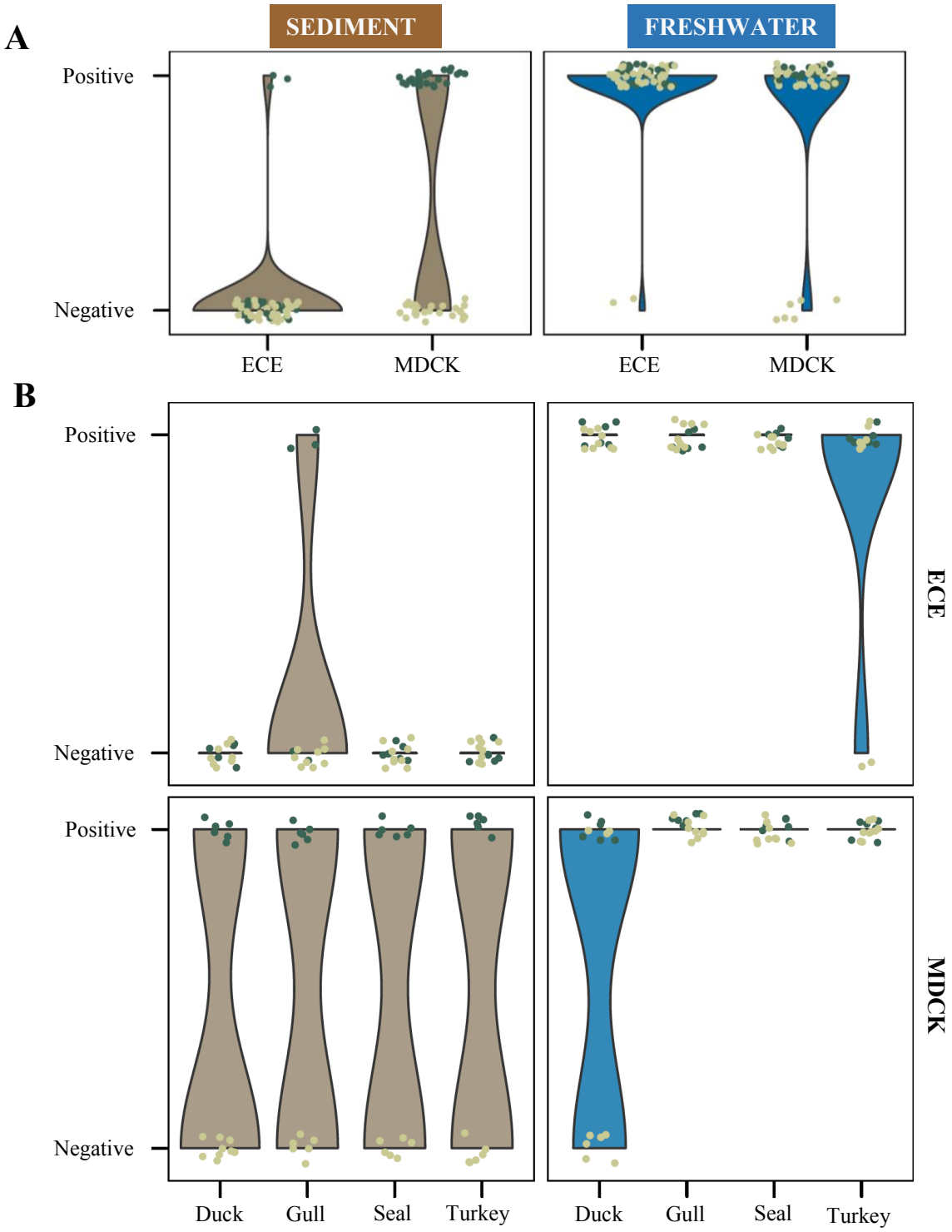


Figure 2: Recovery of influenza A viruses from inoculated sediment and freshwater samples with embryonated chicken eggs (ECE) and Madin-Darby Canine Kidney (MDCK) cells. Graphs show all raw data points including each of the replicate and its distribution separated [A] only by sample type (water vs. sediment) and method (ECE vs. MDCK) or [B] additionally separated by strain. Dark green points represent samples with initial concentration ≥ 100 PFU and light green < 100 PFU. Horizontal lines signify that all samples were either positive or negative. Sediment samples results are shown in the left panels and freshwater samples in the right panels. Influenza A strains are indicated on the x-axis.

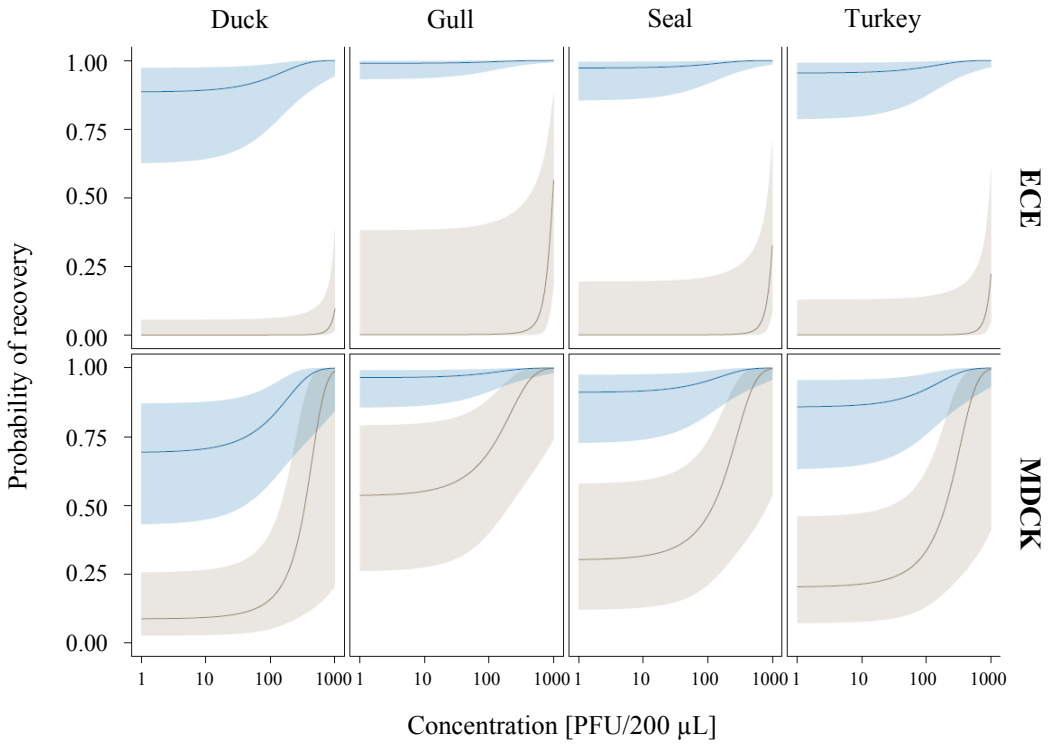


Figure 3: Predicted probability of recovering different influenza A virus strains from inoculated sediment (brown) and freshwater (blue) samples by embryonated chicken eggs (ECE) and Madin-Darby Canine Kidney (MDCK) cells. The x-axis is in log scale.

DISCUSSION

Influenza A virus sequences could be detected by PCR from environmental samples but isolation of IAV failed from those samples (Supplementary **Table S1**). This result did not determine whether the detected virus remained infectious, nor indicate which viral strain was detected. Therefore, dilution experiments were undertaken to determine the minimal viral concentration for four different IAV strains necessary to successfully cultivate IAV from water and water sediment.

Significant differences ($\chi 2 = 12.077$, $p = 0.007$) in the culturing efficiency among used IAV strains were observed (**Figure 2**, **Figure 3**). The IAV duck strain A/Duck/Alberta/35/76 (H1N1), for instance, grew better in embryonated chicken eggs than in MDCK cells, which is consistent with a higher receptor-binding affinity to avian cell surface receptors (58–60). In

contrast, the seal strain A/Seal/Mass/1/80 (H7N7) did not show better growth in MDCK cells compared to embryonated chicken eggs, which is again consistent with its avian origin (50,62–64). For gull strains (H13 and H16) it is known, for example that they do not replicate well in MDCK cell cultures supplemented with trypsin, but efficiently grow in embryonated chicken eggs (65). However, caution is necessary in the interpretation of these findings here given the few strains tested and experiments performed.

While the embryonated eggs worked better for virus recovery from freshwater samples, MDCK cells were more efficient for the recovery from sediment samples, where culturing attempts in embryonated eggs mostly failed ($\chi 2 = 27.68934$, $p < 0.001$). Although previous studies have shown that virus isolation is more efficient in embryonated chicken eggs than in MDCK cells for both swine and avian IAV (66,67), MDCK cells were more tolerant when using sediment, which may contain many additional contaminants compared to freshwater. For example, humic acids and heavy metals may impair the cultivation in embryonated chicken eggs, but may have a smaller effect on MDCK cells. It is known that substances in soil and sediment often inhibit PCR reactions and/or reduce extraction efficiency (65–67) . Bacteria and fungi could also interfere, but are diminished or at least reduced by sterile filtration through 0.22 µm filters. However, a chemical inhibition of virus infection by any inhibitory compounds cannot be excluded. In addition, negative control experiments with IAV free water and sediment did not exhibit viral, bacterial or fungal growth indicating that inhibition – if present - was not of microbial origin.

The low recovery rate from sediment cultures indicates that either there was inhibition, a disruption of virus particles or that most of the virus remained bound to the sediment and were not successfully transferred into the supernatant which was then used for infection. Rapid and tight attachment to sediment or mineral surfaces has been shown previously for different viruses (44,71,72). Furthermore, it could be demonstrated that sediment can prolong viral survival (45,73,74). IAV persistence has been shown to be highest in lake sediment followed by faeces and duck meat (46). Thus, notwithstanding cultivation difficulties, sediment might be a viral reservoir and a good sample source for measuring IAV diversity and investigating its epidemiology and ecology.

Experimentally IAV diluted freshwater samples demonstrated that minimal viral concentrations (1 PFU/200 µL) are needed to infect embryonated eggs or MDCK cells. For

recovery from sediment the minimal viral concentration is suggested to be ≥ 100 PFU/200 µL and MDCK cells are recommended as cultivation method. Previous work has shown that IAV particles are persistent and infectious in freshwater over an extended time, e.g. over 30 days in non-chlorinated river water at 0°C (3) or 100 days at 17°C with salinity of 0 ppt and pH 8.2 (35). The extended stability and low viral concentration needed to seed infection suggests that lake water and sediments could be involved or could enhance IAV intra- and interspecies transmission.

We conclude that IAV cultivation from water samples requires minimal viral titres. Sediment appears to be a source of IAV as well, but further methodological development is needed to improve cultivation efficiency. Our initial PCR screening of sediment samples from different lakes indicated the presence of IAV in 38.8% of samples (Supplementary **Table S1**). However, the inability to culture virus and the findings of the experimental inoculation suggests that the samples had either a virus concentration below the minimum of 100 PFU/200 µL needed to seed infection or that the particles were degraded and no longer infectious. Although IAV surveillance could benefit from environmental sampling, further methodological development will be required to determine the effect of sample type and length of time between viral shedding and sample collection on the ability to cultivate IAV. The results of the current study suggest that such research is warranted.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

ACKNOWLEDGEMENTS

The authors thank Michael Sachtleben for sampling assistance. This work was supported by the grant (IRG 3 - Water) from the Leibniz Research Alliance "INFECTIONS'21 - Transmission Control of Infections in the 21st Century" of the Leibniz Association, Germany (SAS-2015-FZB-LFV).

REFERENCES

1. Stallknecht DE, Shane SM. Host range of avian influenza virus in free-living birds. *Vet Res Commun*. 1988 Mar;12(2–3):125–41.
2. Süss J, Schäfer J, Sinnecker H, Webster RG. Influenza virus subtypes in aquatic birds of eastern Germany. *Arch Virol*. 1994 Mar;135(1–2):101–14.
3. Webster RG, Yakhno M, Hinshaw VS, Bean WJ, Copal Murti K. Intestinal influenza: Replication and characterization of influenza viruses in ducks. *Virology*. 1978 Feb 1;84(2):268–78.
4. Alford RH, Kasel JA, Gerone PJ, Knight V. Human Influenza resulting from aerosol inhalation. *Exp Biol Med*. 1966 Jul 1;122(3):800–4.
5. Fraňková DV, Jirásek A, Tůmová B. Type A influenza: Postmortem virus isolations from different organs in human lethal cases. *Arch Virol*. 1977 Sep;53(3):265–8.
6. Jou WM, Verhoeyen M, Devos R, Saman E, Fang R, Huylebroeck D, et al. Complete structure of the hemagglutinin gene from the human influenza A/Victoria/3/75 (H3N2) strain as determined from cloned DNA. *Cell*. 1980 Mar 1;19(3):683–96.
7. Scholtissek C, Rohde W, Von Hoyningen V, Rott R. On the origin of the human influenza virus subtypes H2N2 and H3N2. *Virology*. 1978 Jun 1;87(1):13–20.
8. Gerber H, others. Clinical features, sequelae and epidemiology of equine influenza. In: *Proceedings 2nd int Conf Equine Infect Dis*, Paris, 1969 [Internet]. 1970 [cited 2016 Jun 7]. p. 63–80. Available from: http://www.cabdirect.org/abstracts/19702205593.html
9. Thomson G, Mumford J, Spooner P, Burrows R, Powell D. The outbreak of equine influenza in England: January 1976. *Vet Rec.* 1977 May 28;100(22):465–8.
10. Tůmová B. Special Issue on animal and human influenzas Equine influenza — A segment in influenza virus ecology. *Comp Immunol Microbiol Infect Dis*. 1980 Jan 1;3(1):45–59.
11. Hinshaw VS, Bean WJ, Webster RG, Easterday BC. The prevalence of influenza viruses in swine and the antigenic and genetic relatedness of influenza viruses from man and swine. *Virology*. 1978 Jan 1;84(1):51–62.
12. Jung K, Ha Y, Chae C. Pathogenesis of Swine influenza virus subtype H1N2 infection in pigs. J *Comp Pathol.* 2005 Feb;132(2–3):179–84.
13. Marozin S, Gregory V, Cameron K, Bennett M, Valette M, Aymard M, et al. Antigenic and genetic diversity among swine influenza A H1N1 and H1N2 viruses in Europe. *J Gen Virol.* 2002;83(4):735–45.
14. Bodewes R, Bestebroer TM, van der Vries E, Verhagen JH, Herfst S, Koopmans MP, et al. Avian influenza A(H10N7) virus–Associated mass deaths among harbor seals. *Emerg Infect Dis.* 2015 Apr;21(4):720–2.
15. Nielsen O, Clavijo A, Boughen JA. Serologic evidence of influenza A infection in marine mammals of Arctic Canada. *J Wildl Dis*. 2001 Oct 1;37(4):820–5.
16. Starick E, Beer M, Hoffmann B, Staubach C, Werner O, Globig A, et al. Phylogenetic analyses of highly pathogenic avian influenza virus isolates from Germany in 2006 and 2007 suggest at least three separate introductions of H5N1 virus. *Vet Microbiol*. 2008 Apr;128(3–4):243–52.
17. Weber S, Harder T, Starick E, Beer M, Werner O, Hoffmann B, et al. Molecular analysis of highly pathogenic avian influenza virus of subtype H5N1 isolated from wild birds and mammals in northern Germany. *J Gen Virol*. 2007 Feb 1;88(2):554–8.
18. Conraths FJ, Sauter-Louis C, Globig A, Dietze K, Pannwitz G, Albrecht K, et al. Highly pathogenic Avian influenza H5N8 in Germany: Outbreak investigations. *Transbound Emerg Dis*. 2016 Feb 1;63(1):10–3.
19. Ellis TM, Bousfield RB, Bissett LA, Dyrting KC, Luk GSM, Tsim ST, et al. Investigation of outbreaks of highly pathogenic H5N1 avian influenza in waterfowl and wild birds in Hong Kong in late 2002. *Avian Pathol.* 2004 Oct 1;33(5):492–505.

20. Koopmans M, Wilbrink B, Conyn M, Natrop G, van der Nat H, Vennema H, et al. Transmission of H7N7 avian influenza A virus to human beings during a large outbreak in commercial poultry farms in the Netherlands. *The Lancet*. 2004 Feb 21;363(9409):587–93.
21. Nili H, Asasi K. Avian influenza (H9N2) outbreak in Iran. *Avian Dis*. 2003 Sep 1;47(s3):828–31.
22. Sturm-Ramirez KM, Ellis T, Bousfield B, Bissett L, Dyrting K, Rehg JE, et al. Reemerging H5N1 influenza viruses in Hong Kong in 2002 are highly pathogenic to ducks. *J Virol*. 2004 May;78(9):4892–901.
23. Chotpitayasunondh T, Ungchusak K, Hanshaoworakul W, Chunsuthiwat S, Sawanpanyalert P, Kijphati R, et al. Human disease from influenza A (H5N1), Thailand, 2004. *Emerg Infect Dis*. 2005 Feb;11(2):201–9.
24. Dinh P, Long H, Tien N, Hien N, Mai L, Phong L, et al. Risk factors for human infection with avian influenza A H5N1, Vietnam, 2004. *Emerg Infect Dis.* 2006;12(12):1841–7.
25. Fouchier RAM, Schneeberger PM, Rozendaal FW, Broekman JM, Kemink SAG, Munster V, et al. Avian influenza A virus (H7N7) associated with human conjunctivitis and a fatal case of acute respiratory distress syndrome. *Proc Natl Acad Sci USA*. 2004 Feb 3;101(5):1356–61.
26. Hien TT, Liem NT, Dung NT, San LT, Mai PP, Chau N van V, et al. Avian influenza A (H5N1) in 10 patients in Vietnam. *N Engl J Med*. 2004 Mar 18;350(12):1179–88.
27. Hinshaw VS, Webster RG, Easterday BC, Bean WJJ. Replication of avian influenza A viruses in mammals. *Infect Immun*. 1981 Nov;34(2):354–61.
28. Murphy BR, Hinshaw VS, Sly DL, London WT, Hosier NT, Wood FT, et al. Virulence of avian influenza A viruses for squirrel monkeys. *Infect Immun*. 1982 Sep;37(3):1119–26.
29. Peiris JSM, Yu WC, Leung CW, Cheung CY, Ng WF, Nicholls JM, et al. Re-emergence of fatal human influenza A subtype H5N1 disease. *Lancet Lond Engl*. 2004 Feb 21;363(9409):617–9.
30. Pensaert M, Ottis K, Vandeputte J, Kaplan MM, Bachmann PA. Evidence for the natural transmission of influenza A virus from wild ducts to swine and its potential importance for man. *Bull World Health Organ*. 1981;59(1):75–8.
31. Subbarao K, Klimov A, Katz J, Regnery H, Lim W, Hall H, et al. Characterization of an avian influenza A (H5N1) virus isolated from a child with a fatal respiratory illness. *Science*. 1998 Jan 16;279(5349):393–6.
32. Ito T, Okazaki K, Kawaoka Y, Takada A, Webster RG, Kida H. Perpetuation of influenza A viruses in Alaskan waterfowl reservoirs. *Arch Virol*. 1995;140(7):1163–72.
33. Markwell DD, Shortridge KF. Possible waterborne transmission and maintenance of influenza viruses in domestic ducks. *Appl Environ Microbiol*. 1982 Jan;43(1):110–5.
34. Brown JD, Swayne DE, Cooper RJ, Burns RE, Stallknecht DE. Persistence of H5 and H7 avian influenza viruses in water. *Avian Dis*. 2007 Mar;51(1 Suppl):285–9.
35. Stallknecht DE, Kearney MT, Shane SM, Zwank PJ. Effects of pH, temperature, and salinity on persistence of avian influenza viruses in water. *Avian Dis*. 1990 Jun;34(2):412–8.
36. Lang AS, Kelly A, Runstadler JA. Prevalence and diversity of avian influenza viruses in environmental reservoirs. *J Gen Virol*. 2008;89(2):509–19.
37. Rohani P, Breban R, Stallknecht DE, Drake JM. Environmental transmission of low pathogenicity avian influenza viruses and its implications for pathogen invasion. *Proc Natl Acad Sci*. 2009 Jun 23;106(25):10365–9.
38. Liu C. Rapid Diagnosis of human influenza infection from nasal smears by means of fluorescein-labeled antibody. *Proc Soc Exp Biol Med*. 1956 Aug 1;92(4):883–7.
39. Mubareka S, Lowen AC, Steel J, Coates AL, García-Sastre A, Palese P. Transmission of influenza virus via aerosols and fomites in the guinea pig model. *J Infect Dis*. 2009 Mar 15;199(6):858–65.
40. Munster VJ, Wit E de, Brand JMA van den, Herfst S, Schrauwen EJA, Bestebroer TM, et al. Pathogenesis and transmission of swine-origin 2009 A(H1N1) influenza virus in ferrets. *Science*. 2009 Jul 24;325(5939):481–3.

41. Roche B, Lebarbenchon C, Gauthier-Clerc M, Chang C-M, Thomas F, Renaud F, et al. Water-borne transmission drives avian influenza dynamics in wild birds: the case of the 2005-2006 epidemics in the Camargue area. *Infect Genet Evol*. 2009 Sep;9(5):800–5.
42. Leung YHC, Zhang L-J, Chow C-K, Tsang C-L, Ng C-F, Wong C-K, et al. Poultry drinking water used for avian influenza surveillance. *Emerg Infect Dis.* 2007 Sep;13(9):1380–2.
43. Dalton MS, Stewart LM, Ip HS. Occurrence of viable avian influenza viruses in water and bed sediments from selected water bodies along the Atlantic Flyway, February and May 2006 and January 2007 [Internet]. *U.S. Geological Survey*; 2009 [cited 2018 Nov 15]. (Open-File Report). Report No.: 2009–1161. Available from: http://pubs.er.usgs.gov/publication/ofr20091161
44. Attinti R, Wei J, Kniel K, Sims JT, Jin Y. Virus' (MS2, ϕX174, and Aichi) attachment on sand measured by atomic force microscopy and their transport through sand columns. *Environ Sci Technol*. 2010 Apr 1;44(7):2426–32.
45. LaBelle RL, Gerba CP. Influence of estuarine sediment on virus survival under field conditions. *Appl Env Microbiol*. 1980 Apr 1;39(4):749–55.
46. Nazir J, Haumacher R, Ike AC, Marschang RE. Persistence of avian influenza viruses in lake sediment, duck feces, and duck meat. *Appl Environ Microbiol*. 2011 May 24;AEM.00415-11.
47. Poulson RL, Luttrell PM, Slusher MJ, Wilcox BR, Niles LJ, Dey AD, et al. Influenza A virus: sampling of the unique shorebird habitat at Delaware Bay, USA. *Open Sci.* 2017 Nov 1;4(11):171420.
48. Khalenkov A, Laver WG, Webster RG. Detection and isolation of H5N1 influenza virus from large volumes of natural water. *J Virol Methods.* 2008 Apr;149(1):180–3.
49. Ward CL, Dempsey MH, Ring CJA, Kempson RE, Zhang L, Gor D, et al. Design and performance testing of quantitative real time PCR assays for influenza A and B viral load measurement. *J Clin Virol*. 2004 Mar;29(3):179–88.
50. Webster RG, Hinshaw VS, Bean WJ, Van Wyke KL, Geraci JR, St. Aubin DJ, et al. Characterization of an influenza A virus from seals. *Virology*. 1981 Sep 1;113(2):712–24.
51. Alexander DJ, Allan WH, Parsons G. Characterisation of influenza viruses isolated from turkeys in Great Britain during 1963--1977. *Res Vet Sci.* 1979 Jan;26(1):17–20.
52. Hinshaw VS, Air GM, Gibbs AJ, Graves L, Prescott B, Karunakaran D. Antigenic and genetic characterization of a novel hemagglutinin subtype of influenza A viruses from gulls. *J Virol*. 1982 Jun 1;42(3):865–72.
53. Hinshaw VS, Webster RG, Turner B. Novel influenza a viruses isolated from Canadian feral ducks: Including strains antigenically related to swine influenza (Hsw1N1) viruses. *J Gen Virol*. 1978;41(1):115–27.
54. Hirst GK. The quantitative determination of influenza virus and antibodies by means of red cell agglutination. *J Exp Med.* 1942 Jan 1;75(1):49–64.
55. Gaush CR, Smith TF. Replication and plaque assay of influenza virus in an established line of canine kidney cells. *Appl Env Microbiol*. 1968 Apr 1;16(4):588–94.
56. R Core Team. R: A language and environment for statistical computing. *R Found Stat Comput Vienna Austria* [Internet]. Available from: https://www.R-project.org/
57. Lee Y, Nelder JA, Pawitan Y, Nelder JA, Pawitan Y. Generalized linear models with random effects : unified analysis via h-likelihood [Internet]. *Chapman and Hall/CRC*; 2006 [cited 2019 Feb 27]. Available from: https://www.taylorfrancis.com/books/9781420011340
58. Rousset F, Ferdy J-B. Testing environmental and genetic effects in the presence of spatial autocorrelation. *Ecography*. 2014;37(8):781–90.
59. Gambaryan A, Webster R, Matrosovich M. Differences between influenza virus receptors on target cells of duck and chicken. *Arch Virol*. 2002 Jun 1;147(6):1197–208.
60. Kuchipudi SV, Nelli R, White GA, Bain M, Chang KC, Dunham S. Differences in influenza virus receptors in chickens and ducks: Implications for interspecies transmission. *J Mol Genet Med Int J Biomed Res*. 2009 Jan 16;3(1):143–51.

61. Rogers GN, Paulson JC. Receptor determinants of human and animal influenza virus isolates: Differences in receptor specificity of the H3 hemagglutinin based on species of origin. *Virology*. 1983 Jun;127(2):361–73.
62. Geraci, Aubin DS, Barker IK, Webster RG, Hinshaw VS, Bean WJ, et al. Mass mortality of harbor seals: pneumonia associated with influenza A virus. *Science*. 1982 Feb 26;215(4536):1129–31.
63. Lang G, Gagnon A, Geraci JR. Isolation of an influenza A virus from seals. *Arch Virol*. 1981 Sep 1;68(3):189–95.
64. Naeve CW, Webster RG. Sequence of the hemagglutinin gene from influenza virus A/Seal/Mass/1 /80. *Virology*. 1983 Sep 1;129(2):298–308.
65. Keawcharoen J, Spronken MIJ, Vuong O, Bestebroer TM, Munster VJ, Osterhaus ADME, et al. Repository of Eurasian influenza A virus hemagglutinin and neuraminidase reverse genetics vectors and recombinant viruses. *Vaccine*. 2010 Aug 16;28(36):5803–9.
66. Clavijo A, Tresnan DB, Jolie R, Zhou E-M. Comparison of embryonated chicken eggs with MDCK cell culture for the isolation of swine influenza virus. *Can J Vet Res.* 2002 Apr;66(2):117–21.
67. Moresco KA, Stallknecht DE, Swayne DE. Evaluation and attempted optimization of avian embryos and cell culture methods for efficient isolation and propagation of low pathogenicity avian influenza viruses. *Avian Dis*. 2010 Mar 1;54(s1):622–6.
68. Tsai YL, Olson BH. Rapid method for separation of bacterial DNA from humic substances in sediments for polymerase chain reaction. *Appl Env Microbiol*. 1992 Jul 1;58(7):2292–5.
69. Watson RJ, Blackwell B. Purification and characterization of a common soil component which inhibits the polymerase chain reaction. *Can J Microbiol*. 2000 Jul 1;46(7):633–42.
70. Wilson IG. Inhibition and facilitation of nucleic acid amplification. *Appl Environ Microbiol.* 1997 Oct;63(10):3741–51.
71. LaBelle RL, Gerba CP. Influence of pH, salinity, and organic matter on the adsorption of enteric viruses to estuarine sediment. *Appl Env Microbiol*. 1979 Jul 1;38(1):93–101.
72. Loveland J., Ryan J., Amy G., Harvey R. The reversibility of virus attachment to mineral surfaces. *Colloids Surf Physicochem Eng Asp*. 1996 Feb;107:205–21.
73. Chung H, Sobsey MD. Comparative Survival of indicator viruses and enteric viruses in seawater and sediment. *Water Sci Technol.* 1993 Feb 1;27(3–4):425–8.
74. Smith EM, Gerba CP, Melnick JL. Role of sediment in the persistence of enteroviruses in the estuarine environment. *Appl Env Microbiol*. 1978 Apr 1;35(4):685–9.

CHAPTER 3: SUPPLEMENTARY MATERIAL

Supplementary Table S1: Detection of influenza A in sediment samples.

Lake	**Stechlin**			**Dagowsee**			**Müggelsee**		**Weißer See**		**Haussee**		
Site	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**1**	**2**	**1**	**2**	**3**
October 2015													
1 cm	-	+	-	-	-	-	-	+	-	‡	+	+	-
3-4 cm	NA	NA	NA	-	NA	-	-	+	NA	+	-	-	+
January 2016													
1 cm	-	-	-	-	-	-	+	-	‡	‡	+	+	+
3-4 cm	-	-	+	NA	NA	+	+	NA	-	-	-	+	+
April 2016													
1 cm	-	-	+	-	-	+	-	+	-	‡	-	-	-
3-4 cm	-	+	-	-	NA	+	-	-	-	+	-	NA	NA

‡ indicates confirmation by sequencing and BLAST search. PCR products were purified and sanger sequenced on a 3130xl Genetic Analyzer using the BigDye® Terminator v1.1 Cycle Sequencing Kit according to the manufacturer's instructions (Applied Biosystems, CA, USA).

The first and 3rd-4th centimetre of sediment was obtained using a plexiglas tube (length 50 cm, Ø 44 mm) and ruler as a sediment corer. Samples were stored at -20°C till extraction. RNA from sediment samples were extracted using the ZR Soil/Fecal RNA MicroPrep™ (ZymoResearch). The RNA was then transcribed in cDNA using the SuperScript™ III Reverse Transcriptase (Invitrogen) and 10µL of extracted RNA. The second DNA strand was synthesized with the Klenow DNA Polymerase I (New England Biolabs). Conventional PCR of a 109 bp fragment of the matrix gene running 35 cycles was performed according to Ward et al., 2004 (doi: 10.1016/S1386-6532(03)00122-7). Only after a reamplification with another 30 cycles using 1 µL of the PCR product the amplified product was visible on a 1.5% agarose gel stained with Midori Green Direct (Biozym). Isolation of IAV from some of the PCR positive sediment samples using embryonated chicken eggs and MDCK cell cultures as described in the current study failed.

GENERAL DISCUSSION

Pathogen screening in the aquatic environment – Challenges and methodological limits

In water, pathogens, particularly the ones that were selected in the current study are strongly diluted and are most probably part of the rare community. Their detection in a 16S rRNA dataset depends on the sequencing depth, so that their presence might not be detected. Thus, the detection of pathogens in huge water sources such as lakes is quite challenging. In addition to the dilution effect and the much higher abundance of other microorganisms, each method has limitations that might lower the detection rate. Therefore, the focus was mainly on sediment samples, as they might contain higher pathogen concentration due to an accumulation and enhanced persistence or protection of the pathogens, e.g. by binding to soil minerals [107–112]. However, sediment samples are complicated to process as they cannot be used directly as sampled, but have to be washed in order to transfer the microorganisms to a liquid phase, e.g. when cell culture or embryonated chicken eggs are used. However, many microorganisms are attached to the sediment particles so that the washing solution might not release them [111, 113–115]. In addition, sediment can contain many inhibitory substances that decrease PCR or cultivation efficiency [116].

In this study**, influenza A viruses** were detected in 38.8% of lake sediment samples by PCR of a short fragment of the matrix gene (109 bp), but the amplification of larger fragments, e.g. of the hemagglutinin gene and the isolation of virus by using embryonated chicken eggs as well as MDCK cells failed (Appendix **Table A2**). The amplification of larger fragments might have failed, because the RNA was damaged during the bead-beating step at the beginning of the extraction [117, 118]. The spiking experiment that is described in '**Chapter 3** - Recovery of influenza A viruses from lake water and sediments by experimental inoculation' revealed that low virus concentration (in our case 1 PFU/200 µL) in freshwater is sufficient to cause an infection, but relatively high virus concentrations (≥ 100 PFU/200 µL) were required to isolate virus from sediment samples. Furthermore, the samples that were first tested by PCR and then used for cultivation were frozen once or twice prior cultivation (due to logistical reasons such as S3 laboratory in another city (Hamburg) and timely ordering of eggs). It is known that freeze-thaw cycles have a negative effect on the infectivity of influenza A virus particles [119] and in the study of Quinlivan et al. (2004) [120] three fresh, unfrozen samples were positive on initial screening but were negative after freezing. In conclusion, the

best chance to isolate influenza A viruses in sediment samples is to use fresh samples and to prepare a culture on MDCK cells instead of embryonated chicken eggs as they seem to be more robust to environmental samples. In addition, MDCK cell cultures are logistically easier to organize and handle than embryonated chicken eggs. A PCR screening on the short matrix gene fragment can help to select positive samples prior cultivation to reduce sample size for cultivation.

Sediment from two lakes (Dagowsee, Weisser See) and wastewater samples were screened for the presence of ***Clostridioides difficile*** by PCR, qPCR and cultivation (Appendix **Table A1**). As shown in '**Chapter 2 -** Genomic analysis of three *Clostridioides difficile* isolates from urban waters sources' we isolated and characterized *C. difficile* from wastewater inflow and the urban lake 'Weisser See'. All samples, except for wastewater, were PCR negative. The fresh sediment sample from 'Weisser See' form which we isolated *C. difficile* was not tested by PCR (Appendix **Table A1**). To the best of our knowledge there are no data published on the prevalence of *C. difficile* in the aquatic environment in Germany. Studies from other countries have shown much higher isolation rates of *C. difficile* from the aquatic environment [68, 69, 94, 121]. Our results, particularly the negative PCR results support the assumption of a low prevalence in the tested samples, even if the isolation rate or diversity of isolates might have been increased by using fresh samples and by investing more time in screening colonies or applying ethanol treatment, which selects for *C. difficile* spores.

In conclusion, our data provided evidence for the presence of pathogens in the investigated water sources emphasizing the need for improving methods to study the pathogens in the aquatic environment. Furthermore, we were able to isolate viable *C. difficile* cells from sediment samples and recovered infectious influenza A viruses from spiked freshwater and sediment samples indicating water bodies as potential reservoirs for those pathogens as they seem to stay infectious.

Wastewater treatment – Blessing and Curse

In urban areas, particularly big cities with high population densities, wastewater treatment is crucial for maintaining aquatic hygiene and reducing waterborne/-transmitted diseases. Our characterization of the bacterial communities in inflow and effluent samples of a WWTP in Berlin, Germany showed that fecal/enteric and most of other potentially harmful bacteria are heavily reduced or decreased to undetectable levels. However, as the sequencing approach mainly detects abundant species, potential pathogens with low abundance might remain unseen. This situation is also true in the current study, where we could isolate *C. difficile* from wastewater and a lake in Berlin, but the 16S rRNA sequence data did not provide evidence for the presence of *C. difficile*. Although low abundance likely represents a low health risk, it has to be taken into account that a threat might also exist by an exchange of harmful genetic elements such as virulence factors or antibiotic resistances between low abundant and abundant groups [122, 123]. It has been demonstrated, for instance, that toxin-negative *C. difficile* strains can become toxin-producers by horizontal gene transfer [122]. Furthermore, initially low abundant species can become abundant when growth conditions are enhanced. WWTPs release the treated wastewater continuously, so that there is a constant contamination of the environment with the effluent. It is known that WWTP effluents influence and change the receiving environment [124–129] - a change that might favour harmful or pathogenic bacteria.

In the current study of the WWTP, phylogenetic analyses revealed some OTUs were closely related to known human pathogens such as *Acinetobacter baumannii*, *Clostridium perfringens*, *Legionella lytica*, *Pseudomonas aeruginosa* and *Yersinia enterocolitica*. The presence of pathogens could be greatly underestimated when using 16S rRNA data only, because it is likely that pathogens are strongly diluted in wastewater samples and masked by other much more abundant bacteria. For instance, as already mentioned we could detect and isolate *C. difficile* from the same samples used for the 16S rRNA study and even detect the *C. difficile* toxin genes via real-time quantitative PCR, but the 16S rRNA dataset did not provide any evidence for the presence of *C. difficile.* Therefore, it might be possible that a pathogen might be not present, rare or low abundant on 16S rRNA sequence level, but can be detected and isolated by cultivation based approaches.

The relatively high abundance of *Legionella* spp. and *Leptospira* spp. in the WWTP of the current study emphasizes the need for extended standard monitoring. While health risk assessments are mainly based on the evaluation of fecal bacteria, cocci and coliforms (e.g. *E. coli*), other potential threats may be underestimated. Furthermore, natural, non-fecal *E. coli* populations have been found in lake soil samples and thus, it is discussed, if *E. coli*, is really suitable as an indicator for fecal contamination [130]. An environment like the WWTP, which provides a large supply of nutrients and organic matter and may facilitate the formation of stable biofilms, might favour many bacteria other than those of fecal origin that can have pathogenic potential. Both *Legionella* and *Leptospira* species are known to form biofilms in which their persistence can be enhanced [131–133]. Our findings indicate a potential risk posed by such bacterial groups and show that the conditions might be beneficial to highly pathogenic species such as *Legionella pneumophila* or *Leptospira interrogans*. Furthermore, a health risk for workers in WWTPs by *Legionella* spp. has been reported in several studies [134–138].

In conclusion, there is a potential health risk associated with underrepresented groups such as *C. difficile* or by water-adapted potentially pathogenic groups such as *Legionella* species. WWTP might favour the growth of some harmful and pathogenic bacteria and influence the natural aquatic environment via their effluents. Furthermore, eukaryotic pathogens and viruses were not investigated and could be worth including in future health risk assessment studies of WWTPs.

Water as reservoir and vector for pathogens

As shown in **Figure 3** there are many routes for pathogens to get into the aquatic environment. For those pathogens that are able to persist or even grow in the water, it can serve as both reservoir and transmission vector. In a world with increasing urbanization and anthropogenic impact on natural environments, aquatic systems experience increasing (selective) pressure, for instance, due to chemical pollution with antibiotics and biocides [139]. Contamination of water bodies increase, which might not only result in higher pathogen prevalence, but also in environmental conditions such as a changed salinity, higher

temperatures and high nutrients that favour the growth of pathogenic and other harmful bacteria such as multi-resistant ones. The release of antibiotic resistant genes and antibiotics into water via WWTPs or agricultural effluents has an impact on natural bacterial communities and might select for antibiotic resistant environmental bacteria [41, 125, 129]. A change or disturbance of natural microbial communities can also lead to ecological consequences and favour the growth of pathogenic microorganisms (dysbiosis) [27, 28, 140–142]. With climate change the water temperatures increases, which has also been shown to favour pathogens such as *Vibrio* spp. resulting in higher infection rates in the Baltic Area [75].

This doctoral thesis focused on the pathogens influenza A viruses and *C. difficile* and found many *Legionella*-related sequences in the WWTP effluent. **Influenza A** sequences were detected in 26 out of 67 (38.8%) sediment samples indicating that sediments of lakes (heavily) used by waterfowl might be a reservoir for influenza A viruses. The virus might accumulate and be relatively protected in the sediment [87–92] and infect individuals after resuspension in the water phase. We were not able to isolate influenza A virus from the PCR positive samples as discussed above, but we could show by spiking sediment and freshwater samples that the virus can be recovered, when present in sufficient concentrations and thus, infection from environmental samples might be realistic. Therefore, there is a need for further studies investigating influenza A viruses in the aquatic environment to understand in detail its ecology and epidemiology.

C. difficile is mainly known as a nosocomial pathogen, but it could be shown that toxigenic strains are also present in the studied aquatic environments [68, 94, 143]. Furthermore, the isolation of living *C. difficile* cells strongly indicates the potential to cause an infection. However, its ecology and epidemiology in the non-hospital environment is still poorly understood and needs further investigation. As our isolates were closely related to human clinical strains, we assume that humans were the source and that the lake was contaminated by a person during recreational activity. This would support our hypothesis that in urban areas with high population densities there is a higher prevalence of human pathogens in aquatic environments due to a higher chance of contamination. The meaning of the findings of the current study need to be further studied in more detail to get more information, with which a health risk assessment can be modelled.

CONCLUDING REMARKS AND FUTURE PERSPECTIVES

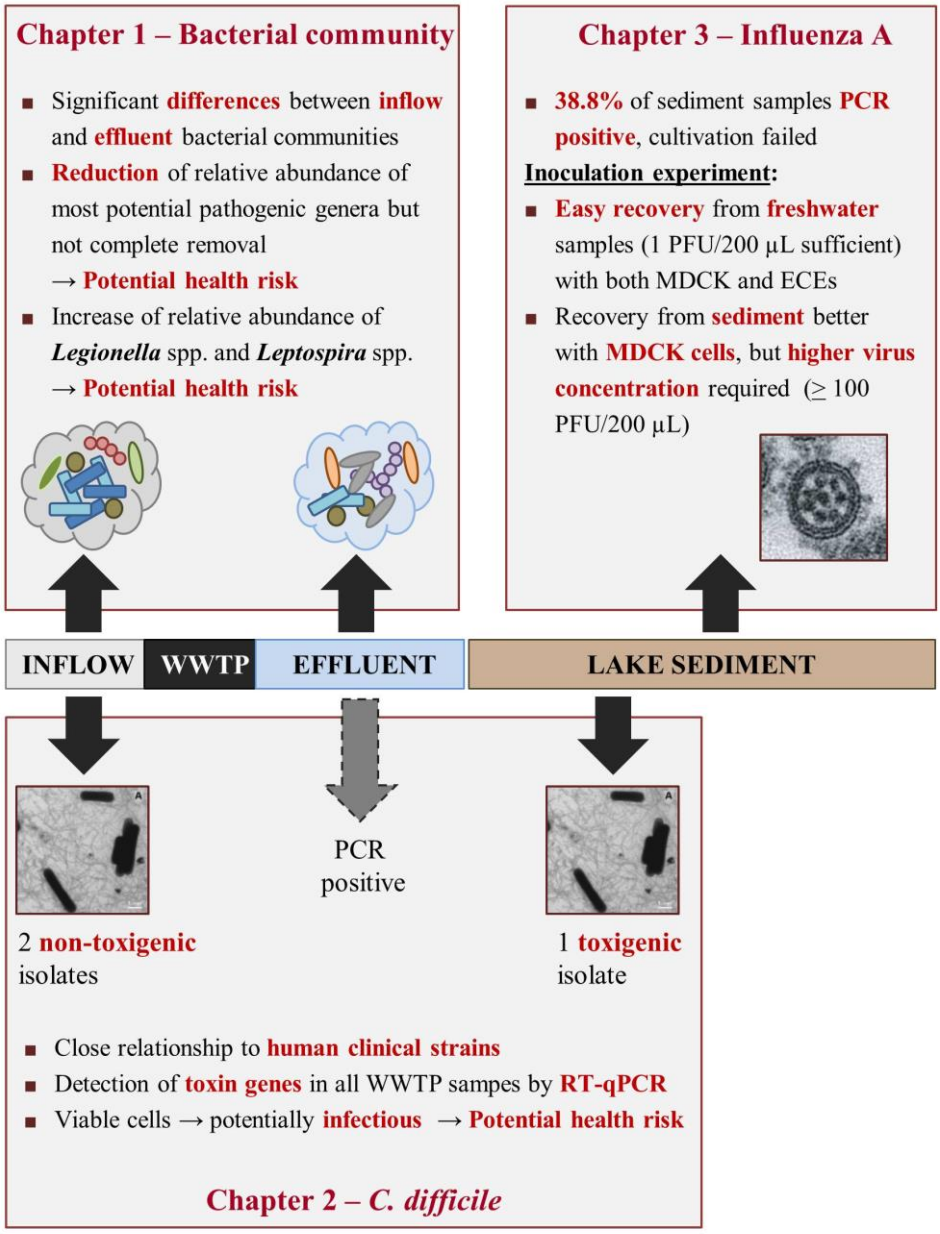


Figure 5: Graphical summary of the most important findings from this doctoral project. Electron microscope images of influenza A virus [86] and *C. difficile* [87] are shown, respectively.

This work encompassed by this doctoral thesis occurred within the Leibniz Alliance Project 'INFECTIONS'21 - Transmission Control of Infections in the 21th Century', a project which characterized bacterial communities of a wastewater treatment plant and examined whether water can act as a hub and/or reservoir for pathogens. **Figure 5** shows a graphical summary of the major findings and in the following the outcome with regard to the initial hypotheses (see H1-H8 under 'Thesis outline', pp. 11-12) is given.

The **bacterial communities** of the inflow and effluent samples of the WWTP differed significantly mainly caused by preferential growth of the phyla *Bacteroidetes*, *Firmicutes*. *Planctomycetes*, *Proteobacteria* and *Verrucomicobia* (**H1**). With the whole 16S rRNA gene sequence information we could relate several OTUs to known species and thus, got a better phylogenetic resolution in comparison to only using the V3-V4 region of the 16S rRNA gene (**H2**). While most of potential pathogens, particularly enteric bacteria, are reduced during the treatmenat, we could define a potential health risk by the presence of potential pathogenic groups such as *Legionella* spp. and other less abundant groups like *Acinetobacter* spp., *Aeromonas* spp. and *Pseudomonas* spp. Phylogenetic analyses revealed high relation of some OTUs to known pathogens such as *Legionella lytica*, *Acinetobacter baumannii* and *Pseudomonas aeruginosa* (**H3**).

We could isolate three ***C. difficile*** strains from urban samples (WWTP and sediment of an urban lake in Berlin, Germany) which were closely related to human strains. This fit to our hypothesis that there is a higher *C. difficile* occurrence in urban than in rural waters, but with having only these few isolates, we cannot draw any conclusions yet on the prevalence of *C. difficile* in the aquatic environment of Germany (**H4**). The isolate from the bathing lake carried the toxin genes, so that a potential health risk exist, even if rather sporadic. Further studies with a higher sampling effort are necessary to test for this hypothesis.

Influenza A viruses was detected in 38.8% of sediment samples by conventional PCR (**H5**), but it was not possible to isolate and grow viruses from those samples. Thus, we cannot draw any conclusions about the diversity (**H6**) or infectivity (**H7**). The inoculation experiment including statistical analyses of cultivation data revealed that the probability of recovery/detection depends on virus concentration, sample type, strain and cultivation method (**H8**). IAV recovery was more effective in embryonated chicken eggs than MDCK cells for freshwater lake dilutions, whereas, MDCK cells were more effective for viral recovery from sediment samples. Low virus concentration (1PFU/200 µL infection volume) was sufficient in most cases to recover IAV from freshwater dilutions. Higher viral concentrations were required for recovery from sediment samples. In conclusion, the inability to virus culture from the lake sediments and the findings of the experimental inoculation suggests that the samples

had either a virus concentration below the minimum of 100 PFU/200 µL needed to seed infection or that the particles were degraded and no longer infectious.

This thesis showed that it is challenging to investigate pathogens in the aquatic environment, but that is also important to increase and improve the use of such environmental samples to understand the ecology and epidemiology of pathogens. This work indicates that water bodies such as wastewater and lake sediments can serve as reservoirs and vectors, even for non-typical water-borne or water-transmitted pathogens such as *C. difficile*. Thus, it is important to keep track on pathogens in aquatic systems and to improve protocols to reliably investigate them in the natural environment including difficult samples such as sediment. Urban areas as pathogen hotspots remain a focus of infectious disease research and need further investigation of microbial communities and pathogens. We need to understand what influence urbanization has on microbial communities and how microbes, particularly pathogens evolve in an urban area. Or was Jean-Jacques Rousseau right by saying "***Cities are the abyss of the human species***" [144]?

Based on the findings of this doctoral thesis, future studies should include the isolation and characterization of *Legionella* spp. from WWTPs as the current study provided strong evidence that there might be a health risk. More spiking experiments as describe in 'CHAPTER 3 - Recovery of influenza A viruses from lake water and sediments by experimental inoculation' are needed to improve protocols for the detection and isolation of pathogens from environmental samples such as water and sediment. Furthermore, studies on *C. difficile* in the aquatic environment of Germany, particularly urban areas, are necessary to further understand its ecology and epidemiology.

> "*If we **pollute** the air, **Water** and soil that keep us alive and well, and destroy the biodiversity that allows natural systems to function, **no amount of money will save us**.* "
>
> David Suzuki

LIST OF REFERENCES

LIST OF REFERENCES

1. Booth DB, Jackson CR. Urbanization of aquatic systems: Degradation Thresholds, stormwater detection, and the limits of mitigation. JAWRA J Am Water Resour Assoc. 1997;33:1077–90.

2. Vörösmarty CJ, Green P, Salisbury J, Lammers RB. Global Water Resources: Vulnerability from climate change and population growth. Science. 2000;289:284–8.

3. Oki T, Kanae S. Global hydrological cycles and world water resources. Science. 2006;313:1068–72.

4. Zhao S, Da L, Tang Z, Fang H, Song K, Fang J. Ecological consequences of rapid urban expansion: Shanghai, China. Front Ecol Environ. 2006;4:341–6.

5. Rockström J, Steffen W, Noone K, Persson Å, Chapin FS, Lambin E, et al. Planetary boundaries: Exploring the safe operating space for humanity. Ecol Soc. 2009;14. https://www.jstor.org/stable/26268316. Accessed 12 Dec 2018.

6. Bogardi JJ, Dudgeon D, Lawford R, Flinkerbusch E, Meyn A, Pahl-Wostl C, et al. Water security for a planet under pressure: interconnected challenges of a changing world call for sustainable solutions. Curr Opin Environ Sustain. 2012;4:35–43.

7. Galli A, Wiedmann T, Ercin E, Knoblauch D, Ewing B, Giljum S. Integrating ecological, carbon and water footprint into a "footprint family" of indicators: Definition and role in tracking human pressure on the planet. Ecol Indic. 2012;16:100–12.

8. Woodward Guy, Perkins Daniel M., Brown Lee E. Climate change and freshwater ecosystems: impacts across multiple levels of organization. Philos Trans R Soc B Biol Sci. 2010;365:2093–106.

9. Lalli CM, Parson TR. Phytoplankton and primary production. Biol Oceanogr Introd. 1997;:40–62.

10. Jumars PA, Penry DL, Baross JA, Perry MJ, Frost BW. Closing the microbial loop: dissolved carbon pathway to heterotrophic bacteria from incomplete ingestion, digestion and absorption in animals. Deep Sea Res Part Oceanogr Res Pap. 1989;36:483–95.

11. Pomeroy LR, leB. Williams PJ, Azam F, Hobbie JE. The microbial loop. Oceanography. 2007;20:28–33.

12. Cole JJ, Findlay S, Pace ML. Bacterial production in fresh and saltwater ecosystems: a cross-system overview. Mar Ecol Prog Ser Oldendorf. 1988;43:1–10.

13. Pfennig N, Widdel F. The bacteria of the sulphur cycle. Phil Trans R Soc Lond B. 1982;298:433–41.

14. Jørgensen BB. Ecology of the bacteria of the sulphur cycle with special reference to anoxic—oxic interface environments. Phil Trans R Soc Lond B. 1982;298:543–61.

15. Francis CA, Beman JM, Kuypers MMM. New processes and players in the nitrogen cycle: the microbial ecology of anaerobic and archaeal ammonia oxidation. ISME J. 2007;1:19–27.

16. Delwiche CC. The nitrogen cycle. Sci Am. 1970;223:136–47.

17. Rigler FH. A tracer study of the phosphorus cycle in lake water. Ecology. 1956;37:550–62.

18. Föllmi KB. The phosphorus cycle, phosphogenesis and marine phosphate-rich deposits. Earth-Sci Rev. 1996;40:55–124.

19. Barsdate RJ, Prentki RT, Fenchel T. Phosphorus cycle of model ecosystems: Significance for decomposer food chains and effect of bacterial grazers. Oikos. 1974;25:239–51.

20. Litchman E. Invisible invaders: non-pathogenic invasive microbes in aquatic and terrestrial ecosystems. Ecol Lett. 2010;13:1560–72.

21. Erb Downward JR, Falkowski NR, Mason KL, Muraglia R, Huffnagle GB. Modulation of post-antibiotic bacterial community reassembly and host response by *Candida albicans*. Sci Rep. 2013;3:2191.

22. Jani AJ, Briggs CJ. The pathogen *Batrachochytrium dendrobatidis* disturbs the frog skin microbiome during a natural epidemic and experimental infection. Proc Natl Acad Sci. 2014;111:E5049–58.

23. Pajares S, Bonilla-Rosso G, Travisano M, Eguiarte LE, Souza V. Mesocosms of aquatic bacterial communities from the Cuatro Cienegas Basin (Mexico): A Tool to test bacterial community response to environmental stress. Microb Ecol. 2012;64:346–58.

24. Castro HF, Classen AT, Austin EE, Norby RJ, Schadt CW. Soil Microbial community responses to multiple experimental climate change drivers. Appl Env Microbiol. 2010;76:999–1007.

25. Lankau RA. Intraspecific variation in allelochemistry determines an invasive species' impact on soil microbial communities. Oecologia. 2011;165:453–63.

26. Ferreyra JA, Ng KM, Sonnenburg JL. The Enteric Two-Step: nutritional strategies of bacterial pathogens within the gut. Cell Microbiol. 2014;16:993–1003.

27. Kamada N, Kim Y-G, Sham HP, Vallance BA, Puente JL, Martens EC, et al. Regulated virulence controls the ability of a pathogen to compete with the gut microbiota. Science. 2012;336:1325–9.

28. van Nood E, Vrieze A, Nieuwdorp M, Fuentes S, Zoetendal EG, de Vos WM, et al. Duodenal infusion of donor feces for recurrent *Clostridium difficile*. N Engl J Med. 2013;368:407–15.

29. Burian SJ, Edwards FG. Historical perspectives of urban drainage. In: Global Solutions for Urban Drainage. 2002. p. 1–16.

30. Mumford L. The city in history: Its origins, its transformations, and its prospects. Houghton Mifflin Harcourt; 1961.

31. Kirby RS. Engineering in history. Courier Corporation; 1990.

32. Webster C. The sewers of Mohenjo-Daro. J Water Pollut Control Fed. 1962;34:116–23.

33. Niemczynowicz J. Urban hydrology and water management – present and future challenges. Urban Water. 1999;1:1–14.

34. Burian SJ, Nix SJ, Pitt RE, Durrans SR. Urban Wastewater Management in the United States: Past, Present, and Future. J Urban Technol. 2000;7:33–62.

35. Asano T, Levine AD. Wastewater reclamation, recycling and reuse: past, present, and future. Water Sci Technol. 1996;33:1–14.

36. Al-Jassim N, Ansari MI, Harb M, Hong P-Y. Removal of bacterial contaminants and antibiotic resistance genes by conventional wastewater treatment processes in Saudi Arabia: Is the treated wastewater safe to reuse for agricultural irrigation? Water Res. 2015;73:277–90.

37. Muga HE, Mihelcic JR. Sustainability of wastewater treatment technologies. J Environ Manage. 2008;88:437–47.

38. Henze M, Harremoes P, Jansen J la C, Arvin E. Wastewater treatment: Biological and chemical processes. Springer Science & Business Media; 2001.

39. Cai L, Zhang T. Detecting human bacterial pathogens in wastewater treatment plants by a high-throughput shotgun sequencing technique. Environ Sci Technol. 2013;47:5433–41.

40. Michael I, Rizzo L, McArdell CS, Manaia CM, Merlin C, Schwartz T, et al. Urban wastewater treatment plants as hotspots for the release of antibiotics in the environment: A review. Water Res. 2013;47:957–95.

41. Rizzo L, Manaia C, Merlin C, Schwartz T, Dagot C, Ploy MC, et al. Urban wastewater treatment plants as hotspots for antibiotic resistant bacteria and genes spread into the environment: A review. Sci Total Environ. 2013;447:345–60.

42. Brody H, Rip MR, Vinten-Johansen P, Paneth N, Rachman S. Map-making and myth-making in Broad Street: the London cholera epidemic, 1854. The Lancet. 2000;356:64–8.

43. Cameron D, Jones IG. John Snow, the Broad Street pump and modern epidemiology. Int J Epidemiol. 1983;12:393–6.

44. Snow J. On the mode of communication of cholera. Edinb Med J. 1856;1:668–70.

45. Koch R. Vierter Bericht des Leiters der deutschen wissenschaftlichen Commission zur Enforschung der Cholera, Geheimen Regierungs Raths Dr Koch. Dtsch Med Wochenschr1884. 1884;10:63.

46. Howard-Jones N. Robert Koch and the cholera vibrio: a centenary. Br Med J Clin Res Ed. 1984;288:379–81.

47. Steyer A, Gutiérrez-Aguirre I, Rački N, Beigot Glaser S, Brajer Humar B, Stražar M, et al. The detection rate of enteric viruses and *Clostridium difficile* in a waste water treatment plant effluent. Food Environ Virol. 2015;7:164–72.

48. Markwell DD, Shortridge KF. Possible waterborne transmission and maintenance of influenza viruses in domestic ducks. Appl Environ Microbiol. 1982;43:110–5.

49. Bal AE, Gravekamp C, Hartskeerl RA, Meza-Brewster JD, Korver H, Terpstra WJ. Detection of leptospires in urine by PCR for early diagnosis of leptospirosis. J Clin Microbiol. 1994;32:1894–8.

50. Babudieri B. Animal reservoirs of leptospires. Ann N Y Acad Sci. 1958;70:393–413.

51. Brockie RE. Leptospiral infections of rodents in the North Island. N Z Vet J. 1977;25:89–96.

52. Schillinger JE, Gannon JJ. Bacterial adsorption and suspended particles in urban stormwater. J Water Pollut Control Fed. 1985;57:384–9.

53. Ward MP. Seasonality of canine leptospirosis in the United States and Canada and its association with rainfall. Prev Vet Med. 2002;56:203–13.

54. Kupek E, De MSSF, De JSP. The relationship between rainfall and human leptospirosis in Florianópolis, Brazil, 1991-1996. Braz J Infect Dis Off Publ Braz Soc Infect Dis. 2000;4:131–4.

55. Desai S, van Treeck U, Lierz M, Espelage W, Zota L, Czerwinski M, et al. Resurgence of field fever in a temperate country: An epidemic of leptospirosis among seasonal strawberry harvesters in Germany in 2007. Clin Infect Dis. 2009;48:691–7.

56. Walters SP, Thebo AL, Boehm AB. Impact of urbanization and agriculture on the occurrence of bacterial pathogens and *stx* genes in coastal waterbodies of central California. Water Res. 2011;45:1752–62.

57. Givens CE, Kolpin DW, Borchardt MA, Duris JW, Moorman TB, Spencer SK. Detection of hepatitis E virus and other livestock-related pathogens in Iowa streams. Sci Total Environ. 2016;566–567:1042–51.

58. Kistemann T, Claßen T, Koch C, Dangendorf F, Fischeder R, Gebel J, et al. Microbial load of drinking water reservoir tributaries during extreme rainfall and runoff. Appl Env Microbiol. 2002;68:2188–97.

59. Plano LR, Garza AC, Shibata T, Elmir SM, Kish J, Sinigalliano CD, et al. Shedding of *Staphylococcus aureus* and methicillin-resistant *Staphylococcus aureus* from adult and pediatric bathers in marine waters. BMC Microbiol. 2011;11:5.

60. Elmir SM, Wright ME, Abdelzaher A, Solo-Gabriele HM, Fleming LE, Miller G, et al. Quantitative evaluation of bacteria released by bathers in a marine water. Water Res. 2007;41:3–10.

61. Stallknecht DE, Kearney MT, Shane SM, Zwank PJ. Effects of pH, temperature, and salinity on persistence of avian influenza viruses in water. Avian Dis. 1990;34:412–8.

62. Brown JD, Swayne DE, Cooper RJ, Burns RE, Stallknecht DE. Persistence of H5 and H7 avian influenza viruses in water. Avian Dis. 2007;51 1 Suppl:285–9.

63. Levin-Edens E, Bonilla N, Meschke JS, Roberts MC. Survival of environmental and clinical strains of methicillin-resistant *Staphylococcus aureus* [MRSA] in marine and fresh waters. Water Res. 2011;45:5681–6.

64. Andre-Fontaine G, Aviat F, Thorin C. Waterborne leptospirosis: Survival and preservation of the virulence of pathogenic *Leptospira* spp. in fresh water. Curr Microbiol. 2015;71:136–42.

65. Hlavsa MC, Cikesh BL, Roberts VA, Kahler AM, Vigar M, Hilborn ED, et al. Outbreaks associated with treated recreational water — United States, 2000–2014. Morb Mortal Wkly Rep. 2018;67:547–51.

66. Kotila SM, Pitkänen T, Brazier J, Eerola E, Jalava J, Kuusi M, et al. *Clostridium difficile* contamination of public tap water distribution system during a waterborne outbreak in Finland. Scand J Public Health. 2013;41:541–5.

67. Reynolds KA. The potential for waterborne spread of MRSA. 2013;:2.

68. Pasquale V, Romano VJ, Rupnik M, Dumontet S, Cižnár I, Aliberti F, et al. Isolation and characterization of *Clostridium difficile* from shellfish and marine environments. Folia Microbiol (Praha). 2011;56:431–7.

69. al Saif N, Brazier JS. The distribution of *Clostridium difficile* in the environment of South Wales. J Med Microbiol. 1996;45:133–7.

70. Levin-Edens E, Soge OO, No D, Stiffarm A, Meschke JS, Roberts MC. Methicillin-resistant *Staphylococcus aureus* from Northwest marine and freshwater recreational beaches. FEMS Microbiol Ecol. 2012;79:412–20.

71. Goodwin KD, McNay M, Cao Y, Ebentier D, Madison M, Griffith JF. A multi-beach study of *Staphylococcus aureus*, MRSA, and enterococci in seawater and beach sand. Water Res. 2012;46:4195–207.

72. Semenza JC, Menne B. Climate change and infectious diseases in Europe. Lancet Infect Dis. 2009;9:365–75.

73. Hunter PR. Climate change and waterborne and vector-borne disease. J Appl Microbiol. 2003;94:37–46.

74. Charron DF, Thomas MK, Waltner-Toews D, Aramini JJ, Edge T, Kent RA, et al. Vulnerability of waterborne diseases to climate change in Canada: A review. J Toxicol Environ Health A. 2004;67:1667–77.

75. Baker-Austin C, Trinanes JA, Taylor NGH, Hartnell R, Siitonen A, Martinez-Urtaza J. Emerging *Vibrio* risk at high latitudes in response to ocean warming. Nat Clim Change. 2013;3:73–7.

76. Viršek MK, Lovšin MN, Koren Š, Kržan A, Peterlin M. Microplastics as a vector for the transport of the bacterial fish pathogen species *Aeromonas salmonicida*. Mar Pollut Bull. 2017;125:301–9.

77. Kirstein IV, Kirmizi S, Wichels A, Garin-Fernandez A, Erler R, Löder M, et al. Dangerous hitchhikers? Evidence for potentially pathogenic *Vibrio* spp. on microplastic particles. Mar Environ Res. 2016;120:1–8.

78. Foley JA, DeFries R, Asner GP, Barford C, Bonan G, Carpenter SR, et al. Global consequences of land use. Science. 2005;309:570–4.

79. Roach WJ, Heffernan JB, Grimm NB, Arrowsmith JR, Eisinger C, Rychener T. Unintended consequences of urbanization for aquatic ecosystems: A case study from the Arizona Desert. BioScience. 2008;58:715–27.

80. Almeida CA, Quintar S, González P, Mallea MA. Influence of urbanization and tourist activities on the water quality of the Potrero de los Funes River (San Luis – Argentina). Environ Monit Assess. 2007;133:459–65.

81. Perkins A, Trimmier M. Recreational Waterborne illnesses: Recognition, treatment, and prevention. Am Fam Physician. 2017;95.

82. Clark RM, Geldreich EE, Fox KR, Rice EW, Johnson CH, Goodrich JA, et al. A waterborne *Salmonella typhimurium* outbreak in Gideon, Missouri: Results from a field investigation. Int J Environ Health Res. 1996;6:187–93.

83. Haley BJ, Cole DJ, Lipp EK. Distribution, diversity, and seasonality of waterborne *Salmonellae* in a rural watershed. Appl Env Microbiol. 2009;75:1248–55.

84. Karagiannis I, Sideroglou T, Gkolfinopoulou K, Tsouri A, Lampousaki D, Velonakis EN, et al. A waterborne *Campylobacter jejuni* outbreak on a Greek island. Epidemiol Infect. 2010;138:1726–34.

85. Andersson Y, De Jong B, Studahl A. Waterborne *Campylobacter* in Sweden: The cost of an outbreak. Water Sci Technol. 1997;35:11.

86. Noda T, Sagara H, Yen A, Takada A, Kida H, Cheng RH, et al. Architecture of ribonucleoprotein complexes in influenza A virus particles. Nature. 2006;439:490–2.

87. Twine SM, Reid CW, Aubry A, McMullin DR, Fulton KM, Austin J, et al. Motility and flagellar glycosylation in *Clostridium difficile*. J Bacteriol. 2009;191:7050–62.

88. Hall IC, O'Toole E. Intestinal flora in new-born infants: With a description of a new pathogenic anaerobe, *Bacillus difficilis*. Am J Dis Child. 1935;49:390–402.

89. Hafiz S, Oakley CL. *Clostridium difficile*: Isolation and characteristics (Plate VIII). J Med Microbiol. 1976;9:129–36.

90. Lyerly DM, Lockwood DE, Richardson SH, Wilkins TD. Biological activities of toxins A and B of *Clostridium difficile*. Infect Immun. 1982;35:1147–50.

91. Lyerly DM, Saum KE, MacDonald DK, Wilkins TD. Effects of *Clostridium difficile* toxins given intragastrically to animals. Infect Immun. 1985;47:349–52.

92. George RH, Symonds JM, Dimock F, Brown JD, Arabi Y, Shinagawa N, et al. Identification of *Clostridium difficile* as a cause of pseudomembranous colitis. Br Med J. 1978;1:695–695.

93. Romano V, Pasquale V, Krovacek K, Mauri F, Demarta A, Dumontet S. Toxigenic *Clostridium difficile* PCR ribotypes from wastewater treatment plants in Southern Switzerland. Appl Environ Microbiol. 2012;78:6643–6.

94. Zidaric V, Beigot S, Lapajne S, Rupnik M. The occurrence and high diversity of *Clostridium difficile* genotypes in rivers. Anaerobe. 2010;16:371–5.

95. Alford RH, Kasel JA, Gerone PJ, Knight V. Human influenza resulting from aerosol inhalation. Exp Biol Med. 1966;122:800–4.

96. Nielsen O, Clavijo A, Boughen JA. Serologic Evidence of influenza A infection in marine mammals of Arctic Canada. J Wildl Dis. 2001;37:820–5.

97. Hinshaw VS, Bean WJ, Webster RG, Easterday BC. The prevalence of influenza viruses in swine and the antigenic and genetic relatedness of influenza viruses from man and swine. Virology. 1978;84:51–62.

98. Starick E, Beer M, Hoffmann B, Staubach C, Werner O, Globig A, et al. Phylogenetic analyses of highly pathogenic avian influenza virus isolates from Germany in 2006 and 2007 suggest at least three separate introductions of H5N1 virus. Vet Microbiol. 2008;128:243–52.

99. Süss J, Schäfer J, Sinnecker H, Webster RG. Influenza virus subtypes in aquatic birds of eastern Germany. Arch Virol. 1994;135:101–14.

100. Weber TP, Stilianakis NI. Inactivation of influenza A viruses in the environment and modes of transmission: A critical review. J Infect. 2008;57:361–73.

101. Weber S, Harder T, Starick E, Beer M, Werner O, Hoffmann B, et al. Molecular analysis of highly pathogenic avian influenza virus of subtype H5N1 isolated from wild birds and mammals in northern Germany. J Gen Virol. 2007;88:554–8.

102. Webster RG, Yakhno M, Hinshaw VS, Bean WJ, Copal Murti K. Intestinal influenza: Replication and characterization of influenza viruses in ducks. Virology. 1978;84:268–78.

103. Ward CL, Dempsey MH, Ring CJA, Kempson RE, Zhang L, Gor D, et al. Design and performance testing of quantitative real time PCR assays for influenza A and B viral load measurement. J Clin Virol. 2004;29:179–88.

104. Lee MS, Chang PC, Shien JH, Cheng MC, Shieh HK. Identification and subtyping of avian influenza viruses by reverse transcription-PCR. J Virol Methods. 2001;97:13–22.

105. Phipps LP, Essen SC, Brown IH. Genetic subtyping of influenza A viruses using RT-PCR with a single set of primers based on conserved sequences within the HA2 coding region. J Virol Methods. 2004;122:119–22.

106. Hoffmann E, Stech J, Guan Y, Webster RG, Perez DR. Universal primer set for the full-length amplification of all influenza A viruses. Arch Virol. 2001;146:2275–89.

107. Skraber S, Schijven J, Italiaander R, Husman AM de R. Accumulation of enteric bacteriophage in fresh water sediments. J Water Health. 2009;7:372–9.

108. Haller L, Poté J, Loizeau J-L, Wildi W. Distribution and survival of faecal indicator bacteria in the sediments of the Bay of Vidy, Lake Geneva, Switzerland. Ecol Indic. 2009;9:540–7.

109. Davies CM, Long JA, Donald M, Ashbolt NJ. Survival of fecal microorganisms in marine and freshwater sediments. Appl Environ Microbiol. 1995;61:1888–96.

110. Nazir J, Haumacher R, Ike AC, Marschang RE. Persistence of avian influenza viruses in lake sediment, duck feces, and duck meat. Appl Environ Microbiol. 2011;:AEM.00415-11.

111. LaBelle RL, Gerba CP. Influence of pH, salinity, and organic matter on the adsorption of enteric viruses to estuarine sediment. Appl Env Microbiol. 1979;38:93–101.

112. LaBelle R, Gerba CP. Investigations into the protective effect of estuarine sediment on virus survival. Water Res. 1982;16:469–78.

113. Attinti R, Wei J, Kniel K, Sims JT, Jin Y. Virus' (MS2, ϕX174, and Aichi) attachment on sand measured by atomic force microscopy and their transport through sand columns. Environ Sci Technol. 2010;44:2426–32.

114. Bai S, Lung W-S. Modeling sediment impact on the transport of fecal bacteria. Water Res. 2005;39:5232–40.

115. LaBelle RL, Gerba CP. Influence of estuarine sediment on virus survival under field conditions. Appl Env Microbiol. 1980;39:749–55.

116. Watson RJ, Blackwell B. Purification and characterization of a common soil component which inhibits the polymerase chain reaction. Can J Microbiol. 2000;46:633–42.

117. Miller DN, Bryant JE, Madsen EL, Ghiorse WC. Evaluation and optimization of DNA extraction and purification procedures for soil and sediment samples. Appl Env Microbiol. 1999;65:4715–24.

118. Bürgmann H, Pesaro M, Widmer F, Zeyer J. A strategy for optimizing quality and quantity of DNA extracted from soil. J Microbiol Methods. 2001;45:7–20.

119. Greiff D, Blumenthal H, Chiga M, Pinkerton H. The effects on biological materials of freezing and drying by vacuum sublimation: II. Effect on influenza virus. J Exp Med. 1954;100:89–101.

120. Quinlivan M, Cullinane A, Nelly M, Maanen K van, Heldens J, Arkins S. Comparison of sensitivities of virus isolation, antigen detection, and nucleic acid amplification for detection of equine influenza virus. J Clin Microbiol. 2004;42:759–63.

121. Simango C. Prevalence of *Clostridium difficile* in the environment in a rural community in Zimbabwe. Trans R Soc Trop Med Hyg. 2006;100:1146–50.

122. Brouwer MSM, Roberts AP, Hussain H, Williams RJ, Allan E, Mullany P. Horizontal gene transfer converts non-toxigenic *Clostridium difficile* strains into toxin producers. Nat Commun. 2013;4:2601.

123. Courvalin P. Transfer of antibiotic resistance genes between gram-positive and gram-negative bacteria. Antimicrob Agents Chemother. 1994;38:1447–51.

124. Carey RO, Migliaccio KW. Contribution of wastewater treatment plant effluents to nutrient dynamics in aquatic systems: A review. Environ Manage. 2009;44:205–17.

125. Gatica J, Cytryn E. Impact of treated wastewater irrigation on antibiotic resistance in the soil microbiome. Environ Sci Pollut Res Int. 2013;20:3529–38.

126. Goudreau SE, Neves RJ, Sheehan RJ. Effects of wastewater treatment plant effluents on freshwater mollusks in the upper Clinch River, Virginia, USA. Hydrobiologia. 1993;252:211–30.

127. Hernando MD, Mezcua M, Fernández-Alba AR, Barceló D. Environmental risk assessment of pharmaceutical residues in wastewater effluents, surface waters and sediments. Talanta. 2006;69:334–42.

128. Wakelin SA, Colloff MJ, Kookana RS. Effect of wastewater treatment plant effluent on microbial function and community structure in the sediment of a freshwater stream with variable seasonal flow. Appl Environ Microbiol. 2008;74:2659–68.

129. Zhang Y, Marrs CF, Simon C, Xi C. Wastewater treatment contributes to selective increase of antibiotic resistance among *Acinetobacter* spp. Sci Total Environ. 2009;407:3702–6.

130. Ishii S, Ksoll WB, Hicks RE, Sadowsky MJ. Presence and growth of naturalized *Escherichia coli* in temperate soils from Lake Superior Watersheds. Appl Env Microbiol. 2006;72:612–21.

131. Declerck P, Behets J, Margineanu A, van Hoef V, De Keersmaecker B, Ollevier F. Replication of *Legionella pneumophila* in biofilms of water distribution pipes. Microbiol Res. 2009;164:593–603.

132. Ristow P, Bourhy P, Kerneis S, Schmitt C, Prevost M-C, Lilenbaum W, et al. Biofilm formation by saprophytic and pathogenic leptospires. Microbiology. 2008;154:1309–17.

133. Murga R, Forster TS, Brown E, Pruckler JM, Fields BS, Donlan RM. Role of biofilms in the survival of *Legionella pneumophila* in a model potable-water system. Microbiology. 2001;147:3121–6.

134. Kusnetsov J, Neuvonen L-K, Korpio T, Uldum SA, Mentula S, Putus T, et al. Two Legionnaires' disease cases associated with industrial waste water treatment plants: A case report. BMC Infect Dis. 2010;10:343.

135. Pascual L, Pérez-Luz S, Amo A, Moreno C, Apraiz D, Catalán V. Detection of *Legionella pneumophila* in bioaerosols by polymerase chain reaction. Can J Microbiol. 2001;47:341–7.

136. Pascual L, Pérez-Luz S, Yáñez MA, Santamaría A, Gibert K, Salgot M, et al. Bioaerosol emission from wastewater treatment plants. Aerobiologia. 2003;19:261–70.

137. Fykse EM, Aarskaug T, Thrane I, Blatny JM. *Legionella* and non-*Legionella* bacteria in a biological treatment plant. Can J Microbiol. 2012;59:102–9.

138. Allestam G, Jong B de, Långmark J. Biological treatment of industrial wastewater: A possible source of Legionella infection. *Legionella*. 2006;:493–6.

139. Taylor NGH, Verner-Jeffreys DW, Baker-Austin C. Aquatic systems: Maintaining, mixing and mobilising antimicrobial resistance? Trends Ecol Evol. 2011;26:278–84.

140. Knights D, Lassen KG, Xavier RJ. Advances in inflammatory bowel disease pathogenesis: linking host genetics and the microbiome. Gut. 2013;62:1505–10.

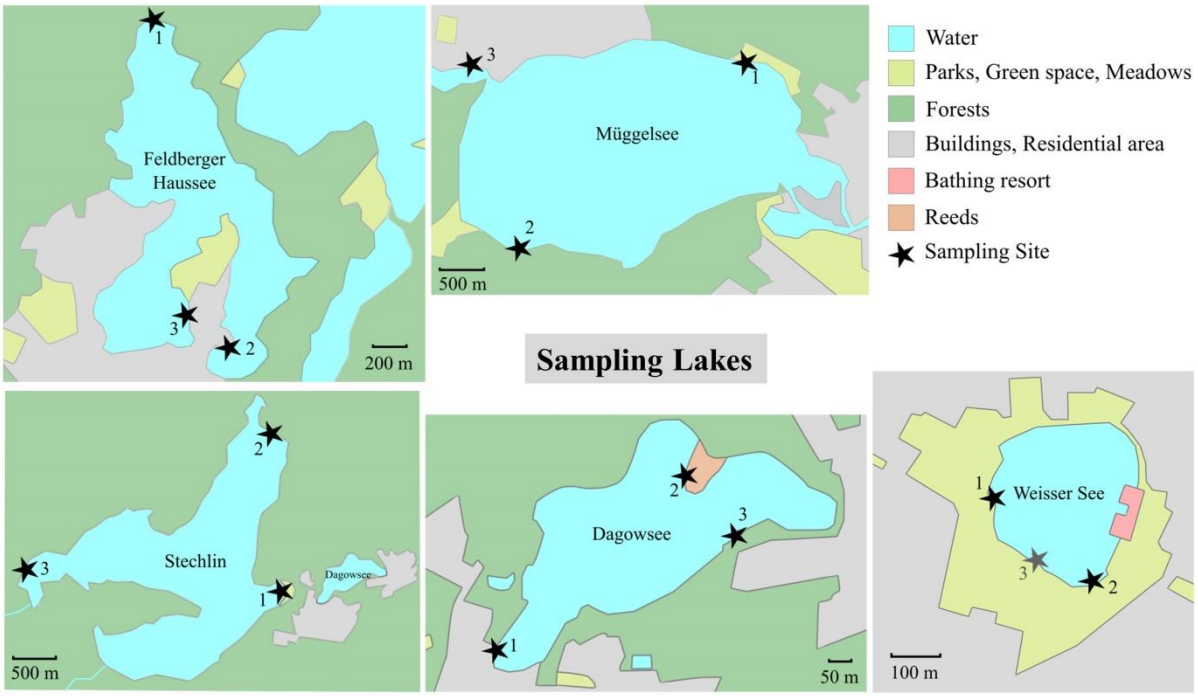
141. Christian N, Whitaker BK, Clay K. Microbiomes: unifying animal and plant systems through the lens of community ecology theory. Front Microbiol. 2015;6. doi:10.3389/fmicb.2015.00869.

142. Sousa WP. Disturbance in Marine Intertidal Boulder Fields: The Nonequilibrium maintenance of species diversity. Ecology. 1979;60:1225–39.

143. Janezic S, Potocnik M, Zidaric V, Rupnik M. Highly Divergent *Clostridium difficile* strains isolated from the environment. PLOS ONE. 2016;11:e0167101.

144. Rousseau J-J. Emile or on education (A. Bloom, Trans.). N Y BasicOriginally Publ 1762. 1979.

APPENDIX



Appendix Figure A1: Lakes in Berlin, Brandenburg and Mecklenburg-Vorpommern, Germany of which surface water and sediment samples were collected every three months during the present PhD project.

Appendix Table A1: Seasonal prevalence of *Clostridioides difficile* in the sediment of two lakes and wastewater revealed by culturing (cult) and PCR detection of the *C. difficile*-specific 16S rRNA sequence (PCR). Da = Dagowsee, WS = Weisser See, WWTP = Wastewater Treatment Plant (+ positive, ‡ positive and selected for genome sequencing, - negative, NA no data available).

Sediment samples from lakes								
	winter		**spring**		**summer**		**autumn**	
	22. Jan		21. Apr		22. Jul		24. Oct	
	PCR	Cult	PCR	Cult	PCR	Cult	PCR	Cult
Da-1	-	-	-	-	-	-	-	-
Da-2	-	-	-	-	-	-	-	-
Da-3	-	-	-	-	-	-	-	-
	winter		**spring**		**summer**		**autumn**	
	05. Feb		14. Apr		30. Jul	15. Aug	18. Oct	
	PCR	Cult	PCR	Cult	PCR	Cult	PCR	Cult
WS-1	-	-	-	-	-	-	-	-
WS-2	-	-	-	-	-	-	-	-
WS-3	NA	NA	NA	NA	-	‡	NA	NA
Water samples from wastewater treatment plant								
	winter		**spring**		**summer**		**autumn**	
	11. Feb		15. Apr		27. Jul		20. Oct	
	PCR	Cult	PCR	Cult	PCR	Cult	PCR	Cult
WWTP_IN	+	-	+	‡	+	‡	+	NA
WWTP_EFF	+	NA	+	NA	-	NA	+	NA

The first centimetre of the sediment was sampled from both lakes using a plexiglas tube (length 50 cm, Ø 44 mm) as a sediment corer. The sediment was placed in 15 mL cryo vials (Carl Roth, Karlsruhe, Germany) and stored at -20°C. DNA from sediment samples was extracted using the NucleoSpin® Soil kit (Macherey Nagel, Düren, Germany) following the manufacturers' instructions. PCRs were perfomed as described in 'Chapter 3 - Genomic analysis of three *Clostridioides difficile* isolates from urban water sources'.

ACKNOWLEDGEMENT

First of all I would like to express my gratitude to everyone who made this PhD project possible and to everyone who helped and supported me.

I would like to start by thanking Prof. Dr. Hans-Peter Grossart and Prof. Alex D. Greenwood, PhD for giving me the opportunity to conduct this research project. I am also very thankful for their encouraging attitude, their guidance and valuable advice. I am very grateful to the entire Leibniz Alliance INFECTIONS'21 for funding (SAS-2015-FZB-LFV), all those interesting scientific discussions, criticism, commendations and advice. Special thanks go to the project coordinators Prof. Dr. Ulrich Schaible and Dr. Susanne Pätzold for their encouraging words, the organization and their patience. In addition, I want to thank my 'co-fighters' Eva Heym, Elisabeth Pfrommer and Dr. Sven Stadtmüller for their scientific and mental support.

I would like to thank the entire Department of Wildlife diseases (IZW) with special thanks to Karin Hönig, Susanne Auls, Adelrose Nekzad, John Galindo, Peter Seeber, Dr. Kristin Mühldorfer, Dr. Gayle McEwen, Saba Mottaghinia, Dr. Matthias Franz and Gábor Czirják. In addition, I am very grateful for the (mainly mental) support and enjoyable time from Silke Ehle, Katrin Hohnstädter, Cornelia Greulich, Beate Peters-Mergner and all PhD students, as well as all members of the workers council.I also want to express my special gratitude to the entire Department of Experimental Limnology (IGB) including Dr. Lars Ganzert, Dr. Luca Zoccarato, Dr. Jason Woodhouse, Uta Mallok, Elke Mach, Michael Sachtleben, Dr. Mina Bizic-Ionescu and Dr. Danny Jonescu. A really big and appreciating 'Thank you' for all the support, help, advice, funny moments and the entertaining fieldwork.

I also would like to thank the HPI and DSMZ for their cooperation and support, particularly Dr. Carola Dreier, Prof. Dr. Gülsah Gabriel, Prof. Dr. Jörg Overmann, Prof. Dr. Ulrich Nübel, Vera Junker and Dr. Thomas Riedel. I greatly appreciate the opportunity to sample the wastewater treatment plant and the 'Tierpark Berlin' and would like to thank Cornelia Prüfer and Dr. Günther Strauss.

Finally, I like to thank my friends and aunts for their support and help, particularly in bad times. THANK YOU VERY MUCH.

STATEMENT OF ORIGINALITY

Erklärung gemäß §12 Abs. 7 der Promotionsordnung zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften an der Mathematisch-Naturwissenschaftlichen Fakultät der Universität Potsdam vom 18. September 2013:

Hiermit versichere ich, dass ich die vorliegende Dissertation selbstständig und ausschließlich mit den angegebenen Mitteln bzw. gekennzeichneten Quellen angefertigt habe und die Arbeit bisher an keiner anderen Hochschule eingereicht worden ist.

The work contained in this thesis has not been previously submitted for a degree at any other higher education institution. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due references are made.

Potsdam, den 11.03.2019 _____

(Daniela Numberger)

CURRICULUM VITAE

Die Seiten 128-130 (Lebenslauf) enthalten persönliche Daten. Sie sind deshalb nicht Bestandteil der Online-Veröffentlichung.

The pages 128-130 (curriculum vitae) contain personal data. Thus, they are not included in the online publication.

Die Seiten 128-130 (Lebenslauf) enthalten persönliche Daten. Sie sind deshalb nicht Bestandteil der Online-Veröffentlichung.

The pages 128-130 (curriculum vitae) contain personal data. Thus, they are not included in the online publication.

Die Seiten 128-130 (Lebenslauf) enthalten persönliche Daten. Sie sind deshalb nicht Bestandteil der Online-Veröffentlichung.

The pages 128-130 (curriculum vitae) contain personal data. Thus, they are not included in the online publication.