Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/jes

JES JOURNAL OF ENVIRONMENTAL SCIENCES

www.jesc.ac.cn

# Epilithic biofilm as a reservoir for functional virulence factors in wastewater-dominant rivers after WWTP upgrade

*Guannan Mao*[1,**], *Jinsong Liang*[2,**], *Qiaojuan Wang*[1,3], *Chen Zhao*[1], *Yaohui Bai*[1,*], *Ruiping Liu*[1,4], *Huijuan Liu*[4], *Jiuhui Qu*[1,4]

[1] *Key Laboratory of Drinking Water Science and Technology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China*

[2] *Harbin Institute of Technology, School of Civil and Environmental Engineering, Shenzhen 518055, China*

[3] *University of Chinese Academy of Sciences, Beijing 100049, China*

[4] *Center for Water and Ecology, State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Tsinghua University, Beijing 100084, China*

ARTICLE INFO

*Article history:*
Received 6 March 2020
Revised 12 May 2020
Accepted 12 May 2020
Available online 21 August 2020

*Keywords:*
Wastewater treatment plant upgrade
Epilithic biofilm
Pathogens
Virulence factors
Health risk

ABSTRACT

Virulence factors (VFs) confer upon pathogens the ability to cause various types of damage or diseases. Wastewater treatment plants (WWTPs) are important point sources for the emission of pathogens and VFs into receiving rivers. Conventional WWTP upgrades are often implemented to improve the water quality of receiving ecosystems. However, knowledge on the pathogens, VFs, and health risks to receiving aquatic ecosystems after upgrade remains limited. In this study, we investigated detailed pathogenic information, including taxa, pathogenicity, and health risk, in two wastewater-dominant rivers after WWTP upgrade. Using 16S rRNA gene sequencing, we screened 14 potential pathogens in water and epilithic biofilm samples, though they were significantly more enriched in the biofilms. Combining 16S rRNA and metagenomic sequencing data, we identified *Pseudomonas* and *Aeromonas* as the dominant pathogenic taxa carrying functional VFs (e.g., mobility and offensive) in the epilithic biofilm. Moreover, strong pathogen-specific VF-host co-occurrence events were observed in the epilithic biofilm samples, indicating the importance of biofilms as reservoirs and vehicles for VFs. Further, we demonstrated that mobility VF is crucial for biofilm formation and pathogens in biofilm carrying offensive VF may be highly invasive. Quantification and health risk assessment suggested that the skin contact risk of *P. aeruginosa* carrying VFs was higher than the acceptable probability of $10^{-4}$ in both water and epilithic biofilm samples, which may threaten ecological and human health.



* Corresponding author.
** These authors contributed equally to this work.
*E-mails:* gnmao@rcees.ac.cn (G. Mao), liangjinsong123@163.com (J. Liang), yhbai@rcees.ac.cn (Y. Bai).

https://doi.org/10.1016/j.jes.2020.05.014
1001-0742/

## Introduction

Bacterial pathogens are usually defined as those with the capacity to cause disease. They also tend to carry single or multiple virulence factors (VFs), which enhance their disease causing potential. Essentially, the ability of bacterial pathogens to establish infection and cause disease (i.e., pathogenicity) can be determined by the number of VFs (Friman et al., 2011; Niu et al., 2013; Sui et al., 2009) and their characterization as either offensive, defensive, motile, or adherent (Diard and Hardt, 2017; Sui et al., 2009; Thomas and Wigneshweraraj, 2014; Wu et al., 2008). VFs are either coded within pathogenicity islands that play key roles in the evolution of bacterial pathogens (Hacker et al., 1997) or within transmissible genetic elements that can spread among bacteria (Schroeder et al., 2017). The rapid emergence and spread of pathogens carrying multiple VFs have been recognized as top health issues by the World Health Organization (Robinson et al., 2016).

Wastewater treatment plants (WWTPs) are important point sources for the emission of pathogens and VFs into receiving rivers (Alexander et al., 2015; Fresia et al., 2019; Rizzo et al., 2013; Ternes et al., 2017) as conventional treatment is usually unable to remove such microbes effectively (Loos et al., 2013; Vilanova et al., 2002). Currently, WWTP upgrades (i.e., additional non-biological treatments) have been implemented in various countries to remove effluent micropollutants and nutrients (Eggen et al., 2014). These additional treatments may decrease pathogen input into receiving rivers, e.g., ozonation is an efficient method used to reduce or inactivate pathogens during tertiary treatment (Alexander et al., 2016; Qiao et al., 2017; Zhuang et al., 2015). However, VFs may be resistant to these supplementary wastewater treatment processes (Wegrzyn and Wegrzyn, 2002; Zhang et al., 2016). Released VFs may enter river water and further into epilithic biofilms (Li et al., 2015). In comparison with water, biofilms with higher bacterial density can increase the probability of VF propagation (Varela and Manaia, 2013). As a result, with long-term development, biofilms can be reservoirs of bacterial pathogens carrying multiple VFs, which is a serious concern for human health (Balcazar et al., 2015).

To date, however, knowledge on pathogens and related VFs in receiving aquatic ecosystems (especially epilithic biofilms) after WWTP upgrade remains poorly understood, which is critical for public health. To address this knowledge gap, we integrated 16S rRNA sequencing, metagenome-assembled genome analysis, and quantitative polymerase chain reaction (qPCR) assay to (1) compare the pathogenic bacterial composition and VFs between water and epilithic biofilm in two receiving rivers after WWTP upgrade; (2) characterize the pathogenic VFs and evaluate their pathogenicity function in water and epilithic biofilms; and, (3) evaluate the potential health risk exposure of target pathogens carrying multiply VFs. The novelty of this work is the focus on epilithic biofilms containing pathogens that harbor pathogenicity genes (i.e., VFs) and specific function and characteristic of VFs in biofilms.

## 1. Materials and methods

### 1.1. *Sampling campaigns*

Epilithic biofilm and water samples were collected from two wastewater-dominant rivers (Tonghui River and Qing River in Beijing, China), which receive wastewater from the Gaobeidian WWTP and Qinghe WWTP, respectively. Both plants were upgraded using ultrafiltration membrane plus ozonation techniques in April 2017 and December 2013 (**Appendix A Table S1**), respectively (Wang et al., 2020). We selected seven sampling sites at each river, one point at effluent outfall and three points upstream and downstream, respectively. The distance between each sampling point was (200 ± 20) m (**Appendix A Fig. S1**). Epilithic biofilms (attached to gravel and stones) are complex matrix-enclosed communities. Epilithic biofilms grow on wet stones in the riverbed. At each site, epilithic biofilm samples were taken along a transect parallel to the river's edge. Cobbles (4-15 cm length) were collected in riffles, which avoided sediment deposition. Details on the two receiving rivers are described in **Appendix A Text S1**.

Water samples were collected in sterile 1000-mL bottles. About 600–800 mL of water was filtered by 0.45-$\mu$m filter membranes (Jinteng, China) in the field followed by the addition of LifeGuard$^{TM}$ preservation solution (MoBio Laboratories Inc., Carlsbad, CA, USA) for further DNA extraction (Liang et al., 2020; Wang et al., 2020; Tang et al., 2016). For biofilm samples, we carefully scraped and collected the substances on stone surfaces at each site using a sterile scraper. The biofilms from each site were transferred to 50-mL sterile tubes and stored at -20°C until DNA extraction. Water samples from Qing River were not collected in December due to the freezing of the river at the seven sites. A total of 105 samples (water: 49; biofilm: 56) were collected during the four seasons in December 2017, March 2018, June 2018, and September 2018.

### 1.2. *DNA extraction and 16S rRNA and shotgun metagenomic sequencing*

Total DNA was extracted from the pellets (for biofilm samples) or from membrane filters (for water samples) using a DNeasy PowerWater Kit (Qiagen, Germany) following the manufacturer's instructions. After detection of DNA density, the amount of DNA in two water samples and 13 biofilm samples was determined to be insufficient for sequencing. Hence, 90 DNA samples (water: 47; biofilm: 43) were sent to the Beijing Genomics Institute (BGI, China) for 16S rRNA sequencing, 16S rRNA sequences of partial water samples were also used in (Wang et al. (2020)) (see **Appendix A Table S2** for details). DNA from 30 samples (water: 14; biofilm: 16) collected in September 2018 was chosen for paired-end shotgun metagenomic sequencing (see **Appendix A Table S3** and **Appendix A Text S2** for details). The sequencing data generated in this study were deposited in the NCBI Sequence Read Archive database under accession number PRJNA418866.

### 1.3. Potential pathogen screening by 16S rRNA sequencing

The raw 16S rRNA sequence reads were initially filtered to remove low-quality reads and barcode primers. The filtered clean reads obtained from 16S rRNA gene amplicon sequencing were analyzed using a QIIME2 pipeline (v2019.4) to infer amplicon sequence variants (ASVs). According to the recommendation of QIIME2 for paired-end ASV analysis, we merged the paired demultiplexed reads and removed low-quality reads based on quality score (phred > 4) (via q2-quality-filter q-score-joined). To screen potential pathogens, we used the normalized sequences obtained in this study to blast the human pathogenic bacteria core dataset downloaded from http://www.mgc.ac.cn/VFs/download.htm (**Appendix A Table S4**) (Chen et al., 2012). The obtained relative abundance tables of pathogenic taxa were exported into the R (v3.4.3) "ggplot2" package.

### 1.4. Identification and characterization of pathogens carrying VFs

#### 1.4.1. Genomic binning

We used genome-centric analysis to identify pathogens carrying VFs from metagenomic datasets. Genomic binning was performed to obtain draft genomes from the water and biofilm metagenomic datasets, respectively (Liang et al., 2020). Briefly, the contigs assembled using MEGAHIT (v1.1.3) were clustered into metagenome assembled genomes (MAGs) using MetaBAT 2 (Kang et al., 2019), MaxBin 2.0 (Wu et al., 2016), and CONCOCT (Alneberg et al., 2014), respectively; the MAGs recovered from the same sample using the three above programs were together refined using the bin_refinement module in the MetaWRAP pipeline to retain only the best representative MAGs (completeness > 50%; contaminant < 10%); dRep was then used to de-replicate (cutoff for filtering: minimum genome completeness 50%; maximum genome contamination 10%) the MAGs from all samples by identifying groups of highly similar MAGs, choosing the best representative for each MAG set, and finally generating a nonredundant MAG set. The coverage of nonredundant MAGs in each sample was estimated using a mapping-based method (BBMap v38.43, under slow mode/high sensitivity) and customized scripts. Abundance of each pathogenic MAG in each sample was calculated based on read mapping percentage and MAG coverage (Liang et al., 2020).

#### 1.4.2. Identification of pathogenic genera carrying VFs

The open reading frames (ORFs) in all contigs of each MAG were predicted using Prodigal (v2.6.3) (Hyatt et al., 2010). To identify potential VF ORFs, the ORFs were aligned against protein sequences in the VFDB_setA_pro.fas file (Liu et al., 2019) downloaded from http://www.mgc.ac.cn/VFs/download.htm using BLASTP (Liang et al., 2020). The identified VFs were classified into two categories based on host (pathogens or non-pathogens): i.e., pathogen-specific VFs and common VFs. In addition, the VFs were classified into offensive, motility, defensive, enzyme, immune suppressive, and unknown based on their functional information. Taxonomic classification of MAGs was performed using GTDB-Tk (v0.3.2) (Parks et al., 2018). The potential pathogens were identified based on the taxonomic list derived from VFDB database, which was the same as taxonomic affiliation of 16S rRNA sequencing. Average nucleotide identity (ANI) among MAGs was calculated using fastANI (v1.1) (Jain et al., 2018) to estimate intergenomic similarity, and MAGs were hierarchically clustered by average linkage.

#### 1.4.3. VFs and pathogen co-occurrence patterns

We conducted network analysis (Ju and Zhang, 2015) between VFs and their hosts to infer associations between pathogen and VF genes in the water and biofilm samples. Correlations among VFs and their hosts were considered statistically significant at $p < 0.05$ and Spearman's $r > 0.6$. All robust correlations identified from pairwise comparison of abundance formed a correlation network. Network analysis data were visualized by Gephi (v0.9.2) (Zhou et al., 2015).

### 1.5. Quantification of target pathogens and risk assessment

*Aeromonas hydrophila* and *Pseudomonas aeruginosa* were targeted because the same genera were identified on the basis of the results of 16S rRNA and metagenomic sequencing. Fecal indicator bacterium (FIB) *Enterococcus* and common enteric pathogen *Salmonella* were also quantified to compare with pathogens carrying VFs. Species-specific functional genes were chosen as biomarkers to quantify fecal indicator bacteria and target pathogenic bacteria. The major adhesin gene *aha*1 (Lee et al., 2008) and out membrane lipoprotein gene *opr*I (DeVos et al., 1997) were chosen as biomarkers to quantify *Aeromonas hydrophila* and *Pseudomonas aeruginosa*, respectively. SYBR Green qPCR assays were performed as described previously (Cui et al., 2019). Details about marker genes, primer sequences, amplification conditions, and standard curve establishment are described in **Appendix A Table S5** and **Appendix A Text S3**.

Microbial risks were further evaluated by quantitative microbial risk assessment (QMRA) (Haas et al., 2014) based on qPCR results (Roser et al., 2015). QMRA was performed for *P. aeruginosa* based on the annual risk model (Rodriguez-Alvarez et al., 2015). A probability of transmission of pathogen to host ($2.0 \times 10^{-6}$) was used (Gerba and Choi, 2006).

$$\text{Annual risk}_{P.aeruginosa} = 1 - \left[1 - \left(\exp\left(-1.05 \times 10^{-4} \times \text{exposure dose}\right)\right)\right]^{365}$$

Gi risk was then determined using exposure doses of *Salmonella* (Boehm et al., 2015).

$$\text{GI risk}_{Salmonella} = 0.17 \times \left(1 - \left(1 + \text{exposure dose}/2884\right)^{-0.3126}\right)$$

### 1.6. Statistical analysis

Beta ($\beta$) diversity was analyzed using principal coordinate analysis (PCoA) based on the weighted UniFrac distance matrix. Analysis of similarity (ANOSIM) was applied to identify differences in the microbial community between water and biofilm samples. The Wilcoxon Paired rank sum test was used

to assess whether pathogen abundance in biofilm samples significantly differed from that in water samples. Statistical analyses were performed using R (v3.4.3).

## 2. Results and discussion

### 2.1. Screening of potential pathogenic genera in water and epilithic biofilm using 16S rRNA sequencing

We compared microbial community composition between the two rivers and between the water and epilithic biofilm samples. The PCoA results (**Appendix A Fig. S2**) indicated that sample type (water vs. biofilm, $R^2 = 0.4543$, ADONIS) was the dominant factor affecting microbial community composition rather than geographical location (Tonghui River vs. Qing River, $R^2 = 0.0451$, ADONIS). In addition, considering that the WWTPs in the two rivers adopted similar upgraded treatment processes, we combined the water or biofilm samples from the two rivers for further pathogenic analysis.

We extracted pathogenic genera information from the water and epilithic biofilm samples, respectively. A total of 14 different genera from two categories, i.e., waterborne/environmental and enteric pathogens, were identified in the water and epilithic biofilm samples, as shown in Fig. 1**a**. Six pathogenic genera showed significantly higher abundance ($p < 0.01$ or $p < 0.001$, Wilcoxon Paired Rank Sum Test) in the epilithic biofilms than in the water samples. Furthermore, a clear separation of pathogenic genera was observed between the epilithic biofilm and water samples based on principal component analysis (PCA) (Fig. 1**b**). The clinical diseases associated with the 14 potential pathogenic genera varied from mild gastroenteritis to severe infection (Chien et al., 2004; Li et al., 2011; Rodriguez-Alvarez et al., 2015) (Fig. 1**a and Appendix A Table S6**). These results indicate that epilithic biofilms may act as reservoirs for clinically relevant pathogens. *Pseudomonas* was the most abundant pathogenic genus in both the epilithic biofilms and water samples. Due to their production of extracellular polymeric substances (EPS), *Pseudomonas* species can resist advanced treatment processes (Alexander et al., 2016), which may explain the dominance of this genus in the epilithic biofilms. *Aeromonas*, another major pathogenic genus in the epilithic biofilms, is a primary cause of several gastrointestinal diseases, ranging from diarrhea to persistent dysentery, and thus poses a risk to public health (Djuikom et al., 2008). The detected enteric pathogenic genera, including *Enterococcus* and *Streptococcus*, showed relative abundances of less than 0.01% in both the water and epilithic biofilm samples, much lower than that reported in urban rivers (0.05% –0.74%) (Sun et al., 2017). This may be because the advanced treatment achieved higher removal efficiencies for the enteric pathogens (Zhang and Farahbakhsh, 2007).

We also found that two bacterial groups (i.e., *Methylotenera* and Rickettsiales) that supported the growth of pathogens were more abundant and prevalent in the epilithic biofilms than in water ($p < 0.001$, Wilcoxon Paired Rank Sum Test; **Appendix A Fig. S3**). Species within the genus *Methylotenera* reportedly use methane as the sole or preferred carbon source and their primary production may generate organic carbon that supports overall growth of opportunistic pathogens (Ling et al., 2016). In addition, Rickettsiales symbionts of free-living amoeba may serve as amplification vehicles for a variety of pathogenic bacteria (Schulz et al., 2016).

Overall, 16S rRNA sequences demonstrated that the epilithic biofilms were enriched in pathogens; however, the pathogenicity of each pathogen needs to be further addressed.

### 2.2. Epilithic biofilms containing pathogens carrying functional VFs revealed by MAGs

To complement the toxicity characteristics of potential pathogens, we further used shotgun metagenomic datasets to extract genomic information in respect to pathogens carrying VFs (Liang et al., 2020; Thompson et al., 2017). As high temperature can promote pathogen growth and increase metabolic activity, September samples were chosen to characterize pathogenic information.

In total, 942 VFs were detected in both water and epilithic biofilm samples, including 614 pathogen-specific VFs and 328 common VFs, accounting for 65.18% and 34.82% of total VFs, respectively. Previous studies have demonstrated many pathogen-specific VFs, especially offensive VFs located within pathogenicity islands, which play keys roles in the evolution of pathogens (Gal-Mor and Finlay, 2006; Hacker and Kaper, 2000; Hentschel and Hacker, 2001). Functional classification also showed that offensive VFs accounted for the highest percentage of VF type in both the epilithic biofilm (44.8% ± 10.1%) and water (55.5% ± 7.6%) samples, indicating that they had more aggressive functions (**Appendix A Fig. S4**), i.e., involved in active invasion of or directly causing damage to the host. Motility VFs also dominated in the epilithic biofilm samples, which assist in the attachment of cells to surfaces and in additional stages of biofilm development (Brown et al., 2019; Shrout et al., 2011).

In total, 513 de-replicated MAGs were recovered from all samples. Among the 135 MAGs carrying the 942 VF genes, 48 MAGs carrying 661 VF genes came from 16 biofilm samples and 87 MAGs carrying 281 VF genes came from 14 water samples. Thus, the biofilm samples demonstrated a higher occurrence of VFs compared to that in the water samples (**Appendix A Table S7**). This can be explained by flagellar cell motility (motility VFs mediated) attachment to abiotic surfaces, and its predominant residence in epilithic biofilms. Pathogen-specific VFs were associated with three dominant pathogenic genera, i.e., *Pseudomonas*, *Aeromonas*, and *Helicobacter* (Fig. 2**a**). To investigate the potential interactions among pathogens and VFs, network analysis of taxa carrying VF co-occurrence patterns was performed (Fig. 2**b and c**). A total of 135 MAGs carrying VFs and six VF types generated 50 nodes and 201 edges in the biofilm network and 42 nodes and 164 edges in the water network based on a significant pairwise correlation ($p < 0.05$, Spearman's $r > 0.6$). In terms of topological properties, the biofilm network presented a higher number of connections per node (average degree = 8.04) compared with the water network (average degree = 7.81) (**Appendix A Table S8**), indicating high connections among the VFs and their hosts in the biofilm network (Balcazar et al., 2015; Faust and Raes, 2012; Newman, 2006). Specifically, *Pseudomonas*, which demonstrated strong co-occurrence with motility VFs in the biofilm network (Spearman's $r = 0.97$), is known to produce

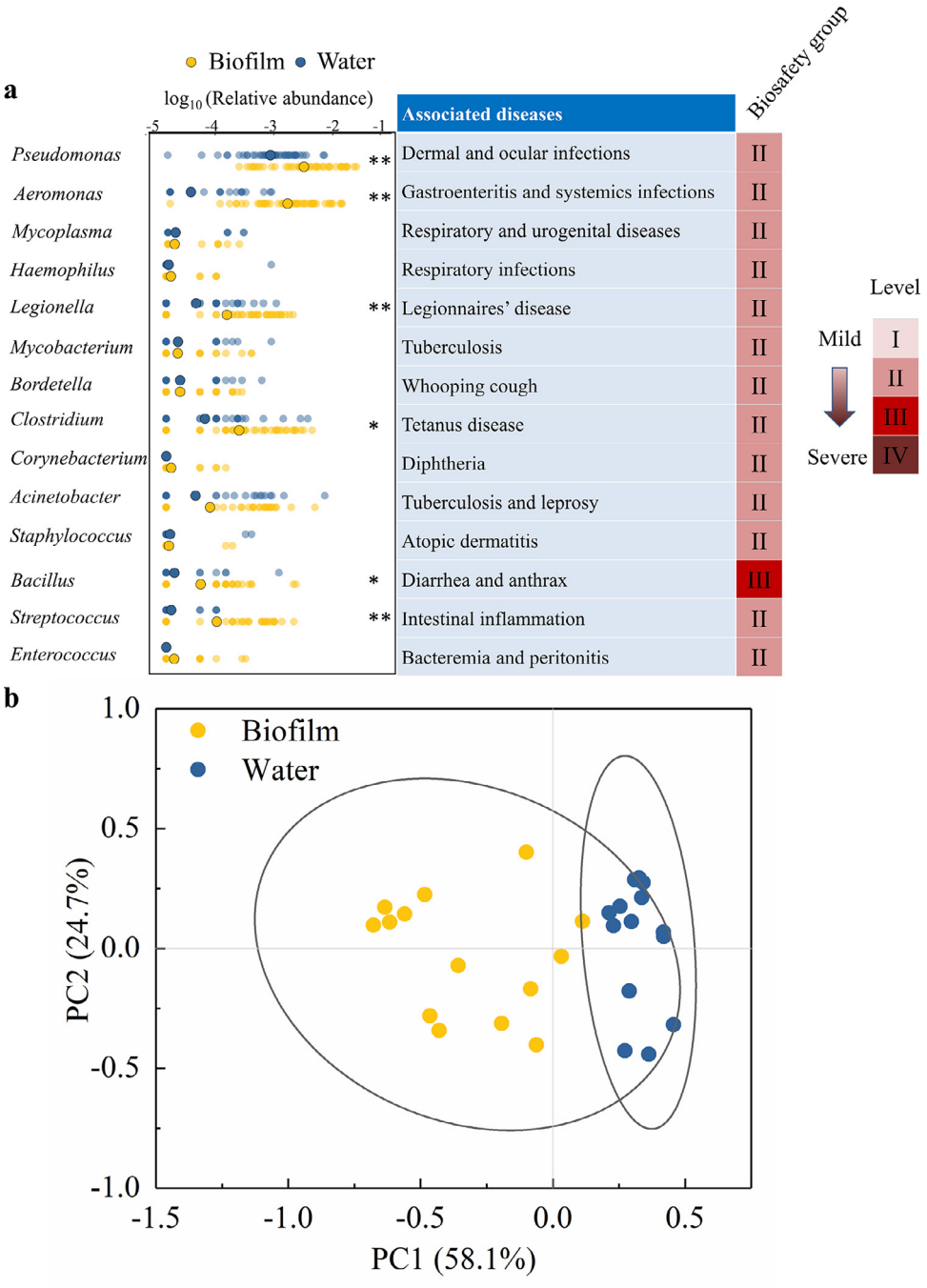


**Fig. 1 – Potential pathogenic genera in water and biofilm samples. Relative abundance of each genus is shown in (a). Correlation coefficient denoted as $p < 0.01$ (*) and $p < 0.001$ (**), with Wilcoxon Paired Rank Sum Test. Right column shows associated diseases (references are listed in Appendix A Table S1) and biosafety group (according to https://my.absa.org/LAI). Principal component analysis (PCA) of 14 potential pathogenic genera indicates a distinct difference between water and biofilm (b).**

motility-secreted virulence mediated by its polar monotrichous flagellum (Li et al., 2015; Rampioni et al., 2009). Swimming and swarming motility are flagellum-dependent behaviors in *Pseudomonas* species. Swarming motility is crucial for biofilm formation, and polar monotrichous flagella aid in motility and colonization of surfaces. Offensive VFs were found to strongly co-occur with *Aeromonas* (Spearman's $r = 0.92$) in the biofilm network, consistent with previous observations that aerolysin in *Aeromonas*, a toxin classified as offensive, plays an integrated and coherent role in the establishment of infections (Sui et al., 2009). These results suggest that *Pseudomonas* carrying motility VFs and *Aeromonas* carrying

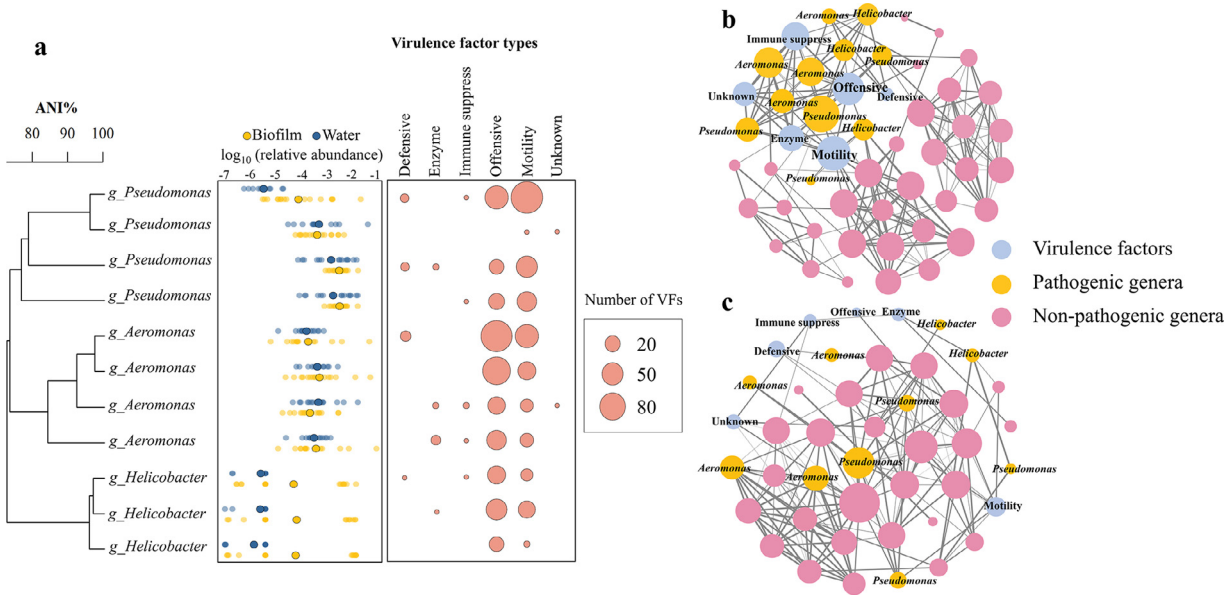


**Fig. 2 – Metagenomic analysis of pathogenic genera carrying Virulence factors (VFs) in water and biofilm samples. Eleven pathogenic metagenome assembled genomes (MAGs) were identified (a). Average nucleotide identity (ANI) clustering shows genome similarity among pathogenic MAGs (far left). Relative abundances of MAGs in water and biofilm samples are shown in middle column. Circles denote mean abundance; dots denote abundance of each MAG; bubbles denote number of VFs. Co-occurrence networks between VFs and pathogenic/non-pathogenic genera in biofilm (b) and water samples (c). Node denotes VF type or MAG carrying VFs and connection (i.e., edge) denotes significant pairwise correlation ($p < 0.05$, Spearman's $r > 0.6$). Network colored by VFs, pathogenic genera, and non-pathogenic genera. Nodes size is proportional to number of connections (i.e., degree) and edge thickness is proportional to Spearman's r.**

offensive VFs were the main pathogenic risks in the epilithic biofilms after WWTP upgrade. It should be noted that, VFs found predominately in pathogens are more directly involved in pathogenicity, i.e., directly cause damage to the host and/or are sufficient to cause disease (Sui et al., 2009). We also highlighted that pathogens identified in this study, no matter by 16S rRNA and metagenomic sequencing, were all potential pathogens.

### 2.3. *Quantification and health risk assessment of target pathogens*

Following the discovery of *Pseudomonas* and *Aeromonas* carrying functional VFs, we investigated whether these two bacteria pose a risk to human health. We first used qPCR to quantify two specific species (i.e., *P. aeruginosa* and *A. hydrophila*) as they are the most clinically relevant pathogens (Djuikom et al., 2008; Pereira et al., 2018). Also, as a reference, fecal indicator bacterium (FIB) *Enterococcus* and common enteric pathogen *Salmonella* were also quantified.

Generally, the concentrations of pathogens were significantly higher ($p < 0.05$, Wilcoxon Paired Rank Sum Test) in the epilithic biofilm samples than in the water samples. The levels of *P. aeruginosa* and *A. hydrophila* were approximately 150 and 60 times greater, respectively, than that of *Salmonella* (Fig. 3a), further indicating that pathogens carrying multiple VFs could be easily enriched or proliferated in wastewater-receiving rivers (Fresia et al., 2019). *Enterococcus* levels were significantly correlated with those of *Salmonella* ($n = 79$, $r = 0.724$, $p < 0.001$), *P. aeruginosa* ($n = 80$, $r = 0.769$, $p < 0.001$), and *A. hydrophila* ($n = 73$, $r = 0.703$, $p < 0.001$) (Fig. 3b), and could serve as an indicator to reflect biomass variation of whole pathogens. The abundance of *P. aeruginosa* showed a relatively strong positive correlation with *A. hydrophila* ($n = 73$, $r = 0.809$, $p < 0.001$) (Fig. 3b), indicating that *P. aeruginosa* and *A. hydrophila* had similar growth conditions or ecological preferences (Hamilton et al., 2017). Both *P. aeruginosa* and *A. hydrophila* are listed as Risk Group II in the ABSA risk group database (https://my.absa.org/LAI). This risk group corresponds to pathogens that can cause mild human disease, as effective treatment and preventive measures are available and risk of infection spread is limited (Varela and Manaia, 2013). The annual risk of infection model associated with recreation in water was used to characterize the health risk of *P. aeruginosa* (Rodriguez-Alvarez et al., 2015). Results showed that the annual risk from *P. aeruginosa* was higher than the acceptable probability of $10^{-4}$ in both the water and epilithic biofilm samples (Fig. 3c). Another health risk scenario, accidental ingestion, was considered for exposure to enteric pathogens, i.e., *Salmonella*. The median ingestion risk in water samples (0.006) was lower than the EPA benchmark of 0.03 (30 in 1000 people), but the median ingestion risk in biofilm samples (0.056) was higher than the acceptable risk benchmark (Fig. 3d). Thus, these results highlight the potential health risks of epilithic biofilm contact in two wastewater-dominant rivers.

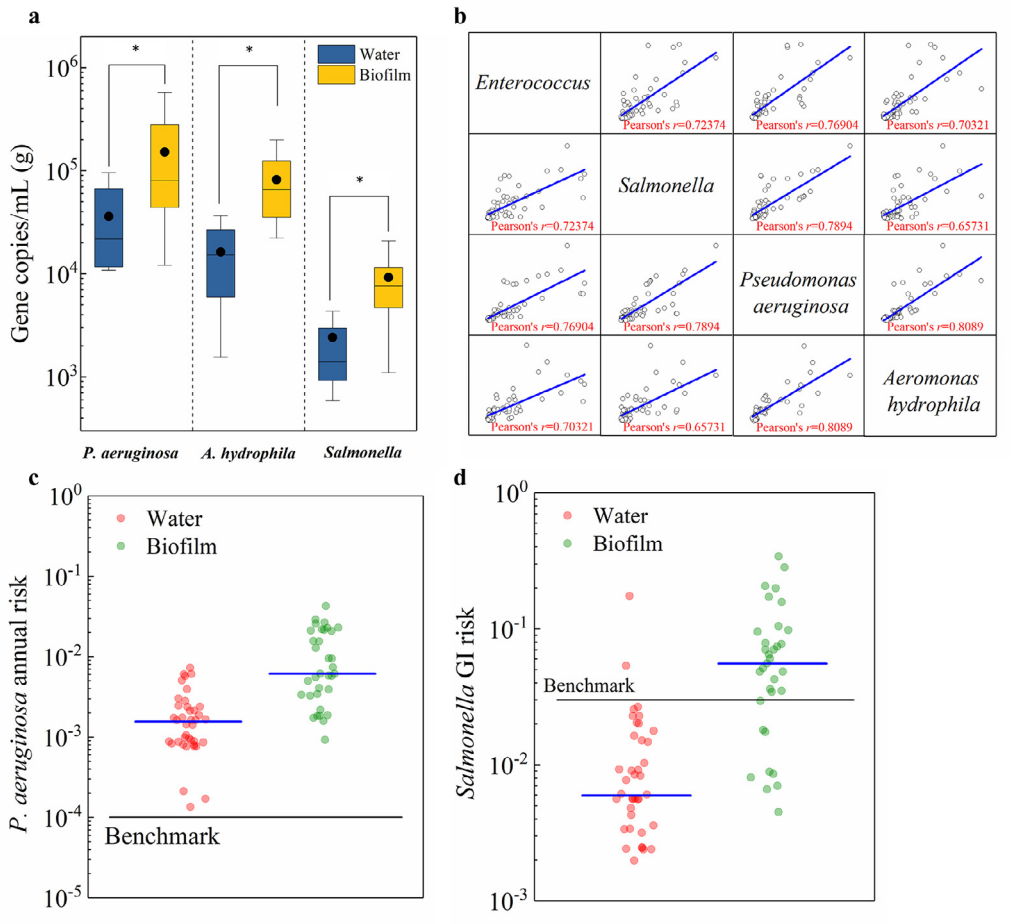


**Fig. 3 – Quantification and risk assessment of target pathogens. Quantification of *P. aeruginosa* and *A. hydrophila* in water and biofilm samples (a). Absolute abundances were determined in gene copies per mL of water or g of biofilm. Correlation coefficient denoted as $p < 0.05$ (*), with Wilcoxon Paired Rank Sum Test. Correlations between enteric pathogens (*Enterococcus* and *Salmonella*) and environmental pathogens (*P. aeruginosa* and *A. hydrophila*) are shown in (b). Pearson's correlation values are superimposed on plots. Probabilities of dermal exposure risk from *P. aeruginosa* (c) and gastrointestinal (GI) risk from *Salmonella* (d) were estimated. Blue lines indicate median. Horizontal lines indicate benchmark of $10^{-4}$ and 0.03 risk levels, respectively.**

## 3. Conclusion

We investigated the characterization and health risk of pathogens in wastewater-dominant rivers after WWTP upgrade. Initial pathogen screening from the 16S rRNA sequences showed 14 human-relevant pathogenic genera in the water and epilithic samples. *Pseudomonas* and *Aeromonas* were the most abundant pathogenic genera and dominated in the epilithic biofilm. Further, we applied the MAG-based approach to identify microbial genomes carrying VFs and demonstrated widespread distribution of VFs in the epilithic biofilm. *Pseudomonas* carrying motility VFs and *Aeromonas* carrying offensive VFs accounted for 31% of total VFs, and thus could pose a risk to public health. Moreover, strong pathogen-specific VF-host co-occurrence events were observed in the epilithic biofilm samples, indicating the importance of epilithic biofilms as reservoirs and vehicles for pathogen-specific VFs. Quantification and health risk assessment suggested high levels of pathogens carrying VF risks in wastewater-receiving rivers after upgrade.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (Nos. 51778603, 51820105011 and 51578537) and Chinese Academy of Sciences (No. QYZDY-SSW-DQC004). The authors thank the Beijing Genomics Institute Central China for providing high-throughput sequencing services.

## Appendix A. Supplementary data

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jes.2020.05.014.

## REFERENCES


Alexander, J., Bollmann, A., Seitz, W., Schwartz, T., 2015. Microbiological characterization of aquatic microbiomes targeting taxonomical marker genes and antibiotic resistance genes of opportunistic bacteria. Sci. Total Environ. 512, 316–325.

Alexander, J., Knopp, G., Dotsch, A., Wieland, A., Schwartz, T., 2016. Ozone treatment of conditioned wastewater selects antibiotic resistance genes, opportunistic bacteria, and induce strong population shifts. Sci. Total Environ. 559, 103–112.

Alneberg, J., Bjarnason, B.S., de Bruijn, I., Schirmer, M., Quick, J., Ijaz, U.Z., et al., 2014. Binning metagenomic contigs by coverage and composition. Nat. Methods 11, 1144–1146.

Balcazar, J.L., Subirats, J., Borrego, C.M., 2015. The role of biofilms as environmental reservoirs of antibiotic resistance. Front Microbiol. 6, 1216.

Boehm, A.B., Soller, J.A., Shanks, O.C., 2015. Human-associated fecal quantitative polymerase chain reaction measurements and simulated risk of gastrointestinal illness in recreational waters contaminated with raw sewage. Environ. Sci. Technol. Lett. 2, 270–275.

Brown, P.C., Borowska, E., Schwartz, T., Horn, H., 2019. Impact of the particulate matter from wastewater discharge on the abundance of antibiotic resistance genes and facultative pathogenic bacteria in downstream river sediments. Sci. Total Environ. 649, 1171–1178.

Chen, L.H., Xiong, Z.H., Sun, L.L., Yang, J., Jin, Q., 2012. VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. Nucleic. Acids Res. 40, D641–D645.

Chien, M.C., Morozova, I., Shi, S.D., Sheng, H.T., Chen, J., Gomez, S.M., et al., 2004. The genomic sequence of the accidental pathogen legionella pneumophila. Science 305, 1966–1968.

Cui, Q.J., Huang, Y., Wang, H., Fang, T.T., 2019. Diversity and abundance of bacterial pathogens in urban rivers impacted by domestic sewage. Environ. Pollut. 249, 24–35.

DeVos, D., Lim, A., Pirnay, J.P., Struelens, M., Vandenvelde, C., Duinslaeger, L., et al., 1997. Direct detection and identification of *Pseudomonas aeruginosa* in clinical samples such as skin biopsy specimens and expectorations by multiplex PCR based on two outer membrane lipoprotein genes. Opri. Oprl. J. Clin. Microbiol. 35 (6), 1295–1299.

Diard, M., Hardt, W.D., 2017. Evolution of bacterial virulence. FEMS Microbiol. Rev. 41 (5), 679–697.

Djuikom, E., Njine, T., Nola, M., Kemka, N., Togouet, S.H.Z., Jugnia, L.B., 2008. Significance and suitability of *Aeromonas hydrophila* vs. fecal coliforms in assessing microbiological water quality. World J. microbiol. Biotechnology 24, 2665–2670.

Eggen, R.I.L., Hollender, J., Joss, A., Scharer, M., Stamm, C., 2014. Reducing the discharge of micropollutants in the aquatic environment: the benefits of upgrading wastewater treatment plants. Environ. Sci. Technol. 48 (14), 7683–7689.

Faust, K., Raes, J., 2012. Microbial interactions: from networks to models. Nat. Rev. Microbiol. 10 (8), 538–550.

Fresia, P., Antelo, V., Salazar, C., Gimenez, M., D'Alessandro, B., Afshinnekoo, E., et al., 2019. Urban metagenomics uncover antibiotic resistance reservoirs in coastal beach and sewage waters. Microbiome 7, 35.

Friman, V.P., Hiltunen, T., Jalasvuori, M., Lindstedt, C., Laanto, E., Ormala, A.M., et al., 2011. High temperature and bacteriophages can indirectly select for bacterial pathogenicity in environmental reservoirs. PLoS One 6 (3), e17651.

Gal-Mor, O., Finlay, B.B., 2006. Pathogenicity islands: a molecular toolbox for bacterial virulence. Cell Microbiol 8 (11), 1707–1719.

Gerba, C.P., Choi, C.Y., 2006. Viruses in foods. In: Food Virology: Past, Present and Future (Goyal, S.M., Ed.). Springer, New York, USA, pp. 1–4.

Haas, C.N., Rose, J.B., Gerba, C.P., 2014. Quantitative Microbial Risk Assessment. John Wiley & Sons, Malden, MA, USA, pp. 979–980.

Hacker, J., BlumOehler, G., Muhldorfer, I., Tschape, H., 1997. Pathogenicity islands of virulent bacteria: Structure, function and impact on microbial evolution. Mol. Microbiol. 23 (6), 1089–1097.

Hacker, J., Kaper, J.B., 2000. Pathogenicity islands and the evolution of microbes. Annu. Rev. Microbiol. 54, 641–679.

Hamilton, K.A., Ahmed, W., Palmer, A., Smith, K., Toze, S., Haase, C.N., 2017. Seasonal assessment of opportunistic premise plumbing pathogens in roof-harvested rainwater tanks. Environ. Sci. Technol. 51 (3), 1742–1753.

Hentschel, U., Hacker, J., 2001. Pathogenicity islands: the tip of the iceberg. Microbes. Infect. 3 (7), 545–548.

Hyatt, D., Chen, G.L., LoCascio, P.F., Land, M.L., Larimer, F.W., Hauser, L.J., 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinform. 11, 119.

Jain, C., Rodriguez-R, L.M., Phillippy, A.M., Konstantinidis, K.T., Aluru, S., 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat. Commun. 9 (1), 5114.

Ju, F., Zhang, T., 2015. Bacterial assembly and temporal dynamics in activated sludge of a full-scale municipal wastewater treatment plant. ISME J. 9 (3), 683–695.

Kang, D.W.D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., et al., 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. Peer J 7, e7359.

Lee, D.Y., Lauder, H., Cruwys, H., Falletta, P., Beaudette, L.A., 2008. Development and application of an oligonucleotide microarray and real-time quantitative PCR for detection of wastewater bacterial pathogens. Sci. Total Environ. 398 (1-3), 203–211.

Li, Y.L., Bai, F., Xia, H.M., Zhuang, L.N., Xu, H.J., Jin, Y.X., et al., 2015. A novel regulator PA5022 (aefA) is involved in swimming motility, biofilm formation and elastase activity of *Pseudomonas aeruginosa*. Microbiol. Res. 176, 14–20.

Li, Y.X., Liu, Y.C., Zhou, Z.M., Huang, H.Q., Ren, Y., Zhang, Y.T., et al., 2011. Complete genome sequence of *Aeromonas veronii* strain B565. J. Bacteriol. 193 (13), 3389–3390.

Liang, J., Mao, G., Yin, X., Ma, L., Liu, L., Bai, Y., et al., 2020. Identification and quantification of bacterial genomes carrying antibiotic resistance genes and virulence factor genes for aquatic microbiological risk assessment. Water Res. 168 (1), 115160.

Ling, F.Q., Hwang, C.A., LeChevallier, M.W., Andersen, G.L., Liu, W.T., 2016. Core-satellite populations and seasonality of water meter biofilms in a metropolitan drinking water distribution system. ISME J. 10 (3), 582–595.

Liu, B., Zheng, D.D., Jin, Q., Chen, L.H., Yang, J., 2019. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. Nucleic. Acids Res. 47 (1), 687–692.

Loos, R., Carvalho, R., Antonio, D.C., Cornero, S., Locoro, G., Tavazzi, S., et al., 2013. EU-wide monitoring survey on emerging polar organic contaminants in wastewater treatment plant effluents. Water Res. 47 (17), 6475–6487.

Newman, M.E.J., 2006. Modularity and community structure in networks. Proc. Natl. Acad. Sci. U.S.A. 103 (23), 8577–8582.

Niu, C., Yu, D., Wang, Y.L., Ren, H.G., Jin, Y., Zhou, W., et al., 2013. Common and pathogen-specific virulence factors are different in function and structure. Virulence 4 (6), 473–482.

Parks, D.H., Chuvochina, M., Waite, D.W., Rinke, C., Skarshewski, A., Chaumeil, P.A., et al., 2018. A standardized

bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat. Biotechnol. 36, 996–1004.
Pereira, R.P.A., Peplies, J., Mushi, D., Brettar, I., Hofle, M.G., 2018. *Pseudomonas*-specific NGS assay provides insight into abundance and dynamics of *Pseudomonas* species including *P. aeruginosa* in a cooling tower. Front. Microbiol. 9 1958.
Qiao, M., Cao, W., Liu, B.C., Bai, Y.H., Qi, W.X., Zhao, X., et al., 2017. Impact of upgrading wastewater treatment plant on the removal of typical methyl, oxygenated, chlorinated and parent polycyclic aromatic hydrocarbons. Sci. Total Environ. 603, 140–147.
Rampioni, G., Schuster, M., Greenberg, E.P., Zennaro, E., Leoni, L., 2009. Contribution of the RsaL global regulator to *Pseudomonas aeruginosa* virulence and biofilm formation. FEMS Microbiol. Lett. 301 (2), 210–217.
Rizzo, L., Manaia, C., Merlin, C., Schwartz, T., Dagot, C., Ploy, M.C., et al., 2013. Urban wastewater treatment plants as hotspots for antibiotic resistant bacteria and genes spread into the environment: a review. Sci. Total Environ. 447, 345–360.
Robinson, T.P., Bu, D.P., Carrique-Mas, J., Fevre, E.M., Gilbert, M., Grace, D., et al., 2016. . Antibiotic resistance is the quintessential one health issue. Trans. Roy. Soc. Trop. Med. Hyg. 110 (7), 377–380.
Rodriguez-Alvarez, M.S., Weir, M.H., Pope, J.M., Seghezzo, L., Rajal, V.B., Salusso, M.M., et al., 2015. Development of a relative risk model for drinking water regulation and design recommendations for a peri urban region of Argentina. Int. J. Hyg. Environ. Health. 218 (7), 627–638.
Roser, D.J., Van Den Akker, B., Boase, S., Haas, C.N., Ashbolt, N.J., Rice, S.A., 2015. Dose-response algorithms for water-borne *Pseudomonas aeruginosa* folliculitis. Epidemiol. Infect. 143 (7), 1524–1537.
Schroeder, M., Brooks, B.D., Brooks, A.E., 2017. The complex relationship between virulence and antibiotic resistance. Genes 8 (1), 39.
Schulz, F., Martijn, J., Wascher, F., Lagkouvardos, I., Kostanjsek, R., Ettema, T.J.G., et al., 2016. A Rickettsiales symbiont of amoebae with ancient features. Environ. Microbiol 18 (8), 2326–2342.
Shrout, J.D., Tolker-Nielsen, T., Givskov, M., Parsek, M.R., 2011. The contribution of cell-cell signaling and motility to bacterial biofilm formation. MRS Bull. 36 (5), 367–373.
Sui, S.J.H., Fedynak, A., Hsiao, W.W.L., Langille, M.G.I., Brinkman, F.S.L., 2009. The association of virulence factors with genomic islands. PLoS One 4 (12), e8094.
Sun, H.H., He, X.W., Ye, L., Zhang, X.X., Wu, B., Ren, H.Q., 2017. Diversity, abundance, and possible sources of fecal bacteria in the Yangtze River. Appl. Microbiol. Biotechnol. 101 (5), 2143–2152.
Tang, J., Bu, Y., Zhang, X.X., Huang, K., He, X., Ye, L., et al., 2016. Metagenomic analysis of bacterial community composition and antibiotic resistance genes in a wastewater treatment plant and its receiving surface water. Ecotoxicol. Environ. Saf. 132, 260–269.
Ternes, T.A., Prasse, C., Eversloh, C.L., Knopp, G., Cornel, P., Schulte-Oehlmann, U., et al., 2017. Integrated evaluation concept to assess the efficacy of advanced wastewater treatment processes for the elimination of micropollutants and pathogens. Environ. Sci. Technol. 51 (1), 308–319.
Thomas, M.S., Wigneshweraraj, S., 2014. Regulation of virulence gene expression. Virulence 5 (8), 832–834.
Thompson, L.R., Sanders, J.G., McDonald, D., Amir, A., Ladau, J., Locey, K.J., et al., 2017. A communal catalogue reveals Earth's multiscale microbial diversity. Nature 551 (7681), 457–463.
Varela, A.R., Manaia, C.M., 2013. Human health implications of clinically relevant bacteria in wastewater habitats. Environ. Sci. Pollut. Res. 20 (6), 3550–3569.
Vilanova, X., Manero, A., Cerda-Cuellar, M., Blanch, A.R., 2002. The effect of a sewage treatment plant effluent on the faecal coliforms and enterococci populations of the reception river waters. J. Appl. Microbiol. 92 (2), 210–214.
Wang, Q., Liang, J., Zhao, C., Bai, Y., Liu, R., Liu, H., et al., 2020. Wastewater treatment plant upgrade induces the receiving river retaining bioavailable nitrogen sources. Environ. Pollut. 263, 114478.
Wegrzyn, G., Wegrzyn, A., 2002. Stress responses and replication of plasmids in bacterial cells. Microb. Cell. Fact. 1, 10.
Wu, H.J., Wang, A.H.J., Jennings, M.P., 2008. Discovery of virulence factors of pathogenic bacteria. Curr. Opin. Chem. Biol. 12 (1), 93–101.
Wu, Y.W., Simmons, B.A., Singer, S.W., 2016. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics 32 (4), 605–607.
Zhang, K., Farahbakhsh, K., 2007. Removal of native coliphages and coliform bacteria from municipal wastewater by various wastewater treatment processes: Implications to water reuse. Water Res. 41 (2), 2816–2824.
Zhang, B., Xia, Y., Wen, X.H., Wang, X.H., Yang, Y.F., Zhou, J.Z., et al., 2016. The composition and spatial patterns of bacterial virulence factors and antibiotic resistance genes in 19 wastewater treatment plants. PLoS One 11 (12), e0167422.
Zhou, J.Z., He, Z.L., Yang, Y.F., Deng, Y., Tringe, S.G., Alvarez-Cohen, L., 2015. High-throughput metagenomic technologies for complex microbial community analysis: open and closed formats. mBio 6, 1–17.
Zhuang, Y., Ren, H.Q., Geng, J.J., Zhang, Y.Y., Zhang, Y., Ding, L.L., et al., 2015. Inactivation of antibiotic resistance genes in municipal wastewater by chlorination, ultraviolet, and ozonation disinfection. Environ. Sci. Pollut. Res. 22 (9), 7037–7044.