

Earth and Environmental Sciences

Special Topic: Emerging Pollution and Emerging Pollutants

Metagenomic insights into the variation of bacterial communities and potential pathogenic bacteria in drinking water treatment and distribution systemsRuiming Jiang¹, Zihan Li¹, Qisheng Li^{1,2}, Yunqi Liu^{1,3}, Yuning Zhu^{1,3}, Zimo Chen^{1,3}, Peng Liu¹, Shuyu Jia^{1,*}, Hongqiang Ren¹ & Xu-Xiang Zhang^{1,*}¹State Key Laboratory of Pollution Control and Resource Reuse, Environmental Health Research Center, School of the Environment, Nanjing University, Nanjing 210023, China;²China Three Gorges Construction Engineering Corporation, Beijing 100048, China;³Nanjing Foreign Language School, Nanjing 210008, China*Corresponding authors: (emails: jiasy@nju.edu.cn (Shuyu Jia); zhangxx@nju.edu.cn (Xu-Xiang Zhang))

Received 21 February 2022; Revised 6 April 2022; Accepted 11 April 2022; Published online 8 June 2022

Abstract: High-throughput sequencing of 16S rRNA gene amplicons was conducted to characterize the changing patterns of bacterial community and potential pathogens in full-scale drinking water treatment and distribution systems. Results showed that Actinobacteria was the predominant phylum in source water, while Proteobacteria dominated after chlorine disinfection and its relative abundance increased from 40.88%±9.45% to 67.86%±27.10%. The genera *Pseudarthrobacter*, *Arenimonas*, and *Limnohabitans* were effectively removed by chlorination, while *Phreatobacter*, *Undibacterium*, *Pseudomonas*, and *Sphingomonas* within the Proteobacteria phylum were greatly enriched after chlorination. Metagenomic analyses revealed the occurrence of 56 species of potential pathogenic bacteria within 17 genera in drinking water, mainly including *Pseudomonas fluorescens* and five mycobacteria species, which were also persistent in tap water samples. The bacteria were found to be involved in various pathways, among which considerable groups were related to human diseases, including infectious diseases and even cancers.

Keywords: drinking water, chlorination, bacterial community, potential pathogenic bacteria, high-throughput sequencing, metagenomics

INTRODUCTION

Microbiological safety is an important and essential requirement for drinking water in the world, especially for potential pathogens in drinking water that have become a global public health issue. Notably, increasing studies have demonstrated the occurrence and prevalence of potential pathogens in drinking water [1,2], suggesting the increased human exposure to potential pathogens and considerable health risk. More seriously, various potential pathogenic bacteria acquire antibiotic resistance genes and show antibiotic resistance. Antimicrobial resistance is common in both sterilized and unsterilized drinking water systems [3].

These potential antibiotic resistant pathogens will increase the difficulty of therapeutic treatment and influence the therapeutic effect.

Sedimentation and filtration, which have been widely used in the purification of drinking water, cannot separate all microbes from source water, usually leading to the regeneration and proliferation of microbes in drinking water pipelines [4]. Disinfection is the most solid barrier to controlling the potential pathogens' growth and ensuring the microbiological safety in the drinking water, which has been proven effective in preventing potential pathogens from spreading to the tap water [2]. Chlorine is usually added to drinking water in the final disinfection step in the waterworks, and a constant concentration of residual chlorine is also necessary for the drinking water distribution system [5–7]. The microbial composition in tap water is usually remarkably different from that in source water and sand filtered water, because chlorination and pipeline transportation has a great impact on bacterial populations [2,8]. However, various bacteria and some potential pathogens possess several adaptive features as their survival mechanism in extreme drinking water conditions, such as chlorine resistance, slow growth, and biofilm formation [9], so that some bacteria and potential pathogens are still discovered in tap water, including *Acinetobacter*, *Burkholderia* and *Pseudomonas* [10–12]. Tap water is an immediate pathway for human exposure to potential pathogenic bacteria, usually via skin exposure or inhalation of aerosols [13], so the bacterial populations involved in human diseases and potential pathogens in drinking water can pose health risks to humans, especially those with immunodeficiency [14]. Moreover, the functional profiles of the bacterial populations especially related to pathogenicity will alter after the structural changes in the bacterial community along with drinking water treatment and distribution. Li *et al.* [15] predicted the function of the biofilm in a drinking water distribution system based on the microbial structure analyses and indicated that an important proportion of the microbial communities was associated with human infectious diseases. Thus, it is urgent and necessary to conduct a comprehensive insight into the bacterial communities and potential pathogenic bacteria in drinking water treatment and distribution systems.

Culturable microbial indicators are often monitored to evaluate the treatment efficiency and the water quality, and the culture-based methods are simple and cost-effective and considered as the gold standard that has been globally applied to monitor pathogenic bacteria [16]. However, no appropriate media are available for the culture of all the target potential pathogens and the methods cannot comprehensively assess the bacterial diversity and function in drinking water [17]. Recently, molecular detection methods such as polymerase chain reaction (PCR)-based methods have been used to explore the occurrence and abundance of certain potential pathogenic bacteria in drinking water [18,19]. However, PCR-based methods usually aim at detecting only known and representative bacterial pathogens due to the limit of validated primers, and it is difficult to comprehensively explore the microorganisms in the environment. Notably, the technologies based on high-throughput DNA sequencing have been proved a promising and powerful tool for examining the complex microbial community and the related function information [20–22]. Among them, 16S rRNA gene amplicon sequencing is a feasible, quick, and cost-efficient approach to successfully investigate the comprehensive bacterial community and directly detect various potential pathogenic bacteria in different environmental media, such as soil [23], surface water [24], activated sludge [25], and drinking water [26].

In this study, 16S rRNA gene amplicon sequencing was used to provide metagenomic insights into the bacterial community structure and potential pathogens in full-scale drinking water treatment and distribution systems, and revealed the main roles of water chlorination and distribution in shaping bacterial community

and potential pathogenic bacteria. The study also sought to examine the bacterial functional profiles along the treatment and distribution systems. This study comprehensively characterized the changing patterns of bacterial community and potential pathogenic bacteria from source water to tap water, especially focusing on the impact of chlorination disinfection. The results may provide valuable reference for the future research on how to effectively control the dissemination of potential pathogens in drinking water and reduce the human health risks.

MATERIALS AND METHODS

Water sample collection and DNA extraction

Four groups of water samples were collected in a waterworks (Nanjing City, Jiangsu Province, China) with supply quantity of 1,200,000 t/d, including 2 L of the source water from the water pumping station I (SW), 4 L of effluent water from the sedimentation tank (ES), 25 L of filtered water from the filtration tank (FW), and 1000 L of chlorine-disinfected water from the clear water tank (DW) (Supplementary Figure S1). Three tap water sampling points (TWA, TWB and TWC) were located along the transportation pipeline, 10, 30 and 40 kilometers away from the waterworks, respectively (Supplementary Figure S1). In order to minimize the experimental error, samples were collected in triplicate from each sampling site (a total of 21 samples) at the same time in June 2017. For pretreatment, 0.22- μm micropore membranes were used to filter SW, ES, and FW samples, while about 1000 L of DW, TWA, TWB, and TWC water were filtered within 24 h to trap bacterial cells using water purifiers due to the low biomass in the samples. The hollow fiber membranes inside the purifiers were subsequently obtained and mixed with 30 mL of purified water. The bacterial cells were separated from the membranes by ultrasound oscillation. The mixture was then centrifuged at 7500 r min^{-1} for 15 min. Total genomic DNA was extracted from the resulting pellets with the FastDNA Soil Kit (MP Biomedicals, CA, USA), and DNA concentration and purity were measured by microspectrophotometry (NanoDropND-2000, NanoDrop Technologies, Wilmington, DE, USA).

High-throughput sequencing

To comprehensively analyze the bacterial community structure and investigate the fate of potential pathogenic bacteria, PCRs were conducted to amplify V3-V4 region (about 460 bp) of the 16S rRNA gene from each sample using the following primers: V3F (5'-ACTCCTACGGGAGGCAGCAG-3') and V4R (5'-TACNVGGGTATCTAATCC-3') [27]. Each PCR reaction system (50 μL) contained 25 μL of 2 \times EasyTaq[®] PCR SuperMix (Transgene, Beijing, China), 40 ng of template DNA, 2 μL of forward and reverse primer (10 $\mu\text{mol L}^{-1}$), and 21 μL of ddH₂O. The thermal changes of PCRs included initial pre-denaturation at 94°C (3 min), 30 amplification cycles of denaturation at 94°C (30 s), annealing at 62°C (30 s) and extension at 70°C (45 s), and final extension at 70°C (7 min) [4]. Before sequencing, the PCR products were purified using TaKaRa MiniBEST DNA Fragment Purification Kit Ver. 4.0 (TaKaRa, Japan). The appropriate amounts of purified PCR products were then mixed and sent to Jiangsu Zhongyijinda Analytical & Testing Co., Ltd. (Yixing, China) for sequencing on a Miseq platform (Illumina, San Diego, CA, USA). The raw high-throughput sequencing data were uploaded to the publicly available database (NCBI Sequence Read Ar-

chive) and the obtained accession number is PRJNA734573.

Bioinformatics analysis

Mothur software was used to further process the raw sequences, including quality controlling, denoising, chimera removal, and 885,690 clean reads were generated in total. To avoid biases between different samples during the comparison, the sequencing reads were normalized to 29,523 for each sample to pick operational taxonomic units (OTU), calculate richness and diversity indices, and annotate the sequences. OTUs were defined at a similarity of 97% using USEARCH implemented in Quantitative Insights into Microbial Ecology (QIIME). For taxonomic classification, the sequences were aligned by Ribosomal Database Project (RDP) classifier against the Silva SSU database (Release123) and the min confidence threshold was 80% [28]. Richness and diversity indices such as the number of OTUs, Chao 1 index, and Shannon index were calculated using Mothur software.

Detection of potential pathogenic bacteria

The species and strains of human pathogenic bacteria (HPB) were summarized by collecting the information from HPB virulence factor database (<http://www.mgc.ac.cn/VFs/>) [29] and other reference databases [30,31]. The HPB 16S rRNA gene sequence database was created by extracting the sequences from the NCBI GenBank (<http://www.ncbi.nlm.nih.gov/>), including 1525 16S rRNA gene sequences assigned to 527 HPB species. The BLASTN program was applied to search the target sequences from each sample against the offline HPB 16S rRNA gene sequence database [32]. The BLAST outputs were further screened using a strict criterion of identity $\geq 99\%$, alignment length ≥ 150 bp, and mismatch ≤ 1 bp [23]. The percentage of the number of screened sequences in the total sequencing reads (29,523 in this study) was recognized as the relative abundance of potential pathogenic bacteria.

Functional composition analysis

In this study, PICRUSt based on KEGG database was used for functional prediction of bacterial communities [33]. Firstly, MOTHUR software was used to convert the 16S rRNA OTU-shared file (with identity of over 97%) into a BIOM format [34]. The obtained OTU table was then uploaded to the online PICRUSt server on the Galaxy (<http://huttenhower.sph.harvard.edu/galaxy>) and the function of the OTUs was predicted following the developer-recommended instructions (<http://picrust.github.io/picrust>). Default three levels of KEGG pathway hierarchy was chosen and hierarchical data was collapsed to a specified level for PICRUSt predictions.

Statistical analyses

The differences of microbial community structure in drinking water before and after chlorination were calculated with linear discriminant analysis effect size (LEfSe) online (<http://huttenhower.sph.harvard.edu/galaxy>), which highlights both biological relevance and statistical significance [35]. In order to explore the

difference of functional community structure among the samples, principal coordinates analysis (PCoA) was conducted using PAST software ver 3.15 (<http://folk.uio.no/ohammer/past/>). Venn diagram, reflecting the number and proportion of shared potential pathogenic bacteria among all the tap water samples, was generated using Venny v2.1.0 [36]. The biodiversity indices and relative abundance of bacterial community were obtained by calculating the average of the triplicate samples and expressed as the mean±standard deviation. The variation among different samples in drinking water was analyzed by SPSS software using ANOVA (a one-way analysis of variance), and $P<0.05$ was considered statistically significant.

RESULTS

Bacterial community shift during drinking water treatment and distribution

Biodiversity of the drinking water samples during treatment and distribution processes was analyzed by calculating the number of OTUs, Chao 1 index, and Shannon index at 3% cutoff. As for the number of OTUs, SW had the highest biodiversity (4179±1189 OTUs), followed by ES (3062±418 OTUs) and FW (3257±891 OTUs), while DW and TWs displayed considerably less richness ((568±19)–(1087±313) OTUs) (Supplementary Figure S2). Notably, the number of OTUs significantly reduced after chlorination ($P<0.05$). This result was also supported by the analyses of Chao 1 index and Shannon index, revealing that biodiversity slightly decreased after sedimentation and sand filtration, then significantly decreased after chlorination ($P<0.05$), and finally remained relatively stable with only small fluctuations in the process of transportation (Supplementary Figures S3 and S4). Biodiversity shift revealed the important role of drinking water treatment and distribution (especially the chlorination) on the species richness variation.

Microbial classification based on 16S rRNA gene sequencing indicated that the microbial composition of both phyla (Figure 1A) and genera (Figure 1B) changed in the process of the drinking water treatment and distribution. The PCoA plot also showed the bacterial community shift at genus level after chlorination (Figure 2). In general, the phyla Actinobacteria and Proteobacteria had the highest abundance in SW, ES, and FW, which accounted for 74.55%–84.13% of the total relative abundance, followed by Bacteroidetes and Acidobacteria with a relative abundance of more than 3% (Figure 1A). Notably, Actinobacteria (from 35.59%±6.63% to 1.69%±0.16%), Acidobacteria (from 5.03%±3.34% to 0.01%±0.002%), and Bacteroidetes (from 5.41%±1.87% to 0.03%±0.005%), were more sensitive to chlorine because chlorination greatly decreased their relative abundance (Figure 1A). For the drinking water samples after chlorination, Proteobacteria was dominant and its relative abundance increased from 40.88%±9.45% to 67.86%±27.10% after chlorine disinfection, and further increased to 82.82%±12.09%, 85.53%±6.71%, and 98.43%±0.47% in TWA, TWB and TWC, respectively, after pipeline distribution (Figure 1A). However, the relative abundance of most phyla reduced to less than 1% after chlorination. Unusually, the relative abundance of Planctomycetes increased from 3.16%±1.12% to 22.00%±16.45% after chlorine disinfection but soon decreased to 7.21%±7.17% after pipeline distribution.

At genus level, HgcI_clade and CL500-29_marine_group had the highest relative abundance in the pre-chlorination samples, but only a few of them remained after chlorination (Figure 1B). Moreover, other predominant genera, including *Pseudarthrobacter*, *Arenimonas* and *Limnohabitans* also obviously decreased after chlorination. On the contrary, the relative abundance of several bacterial genera within the Proteo-

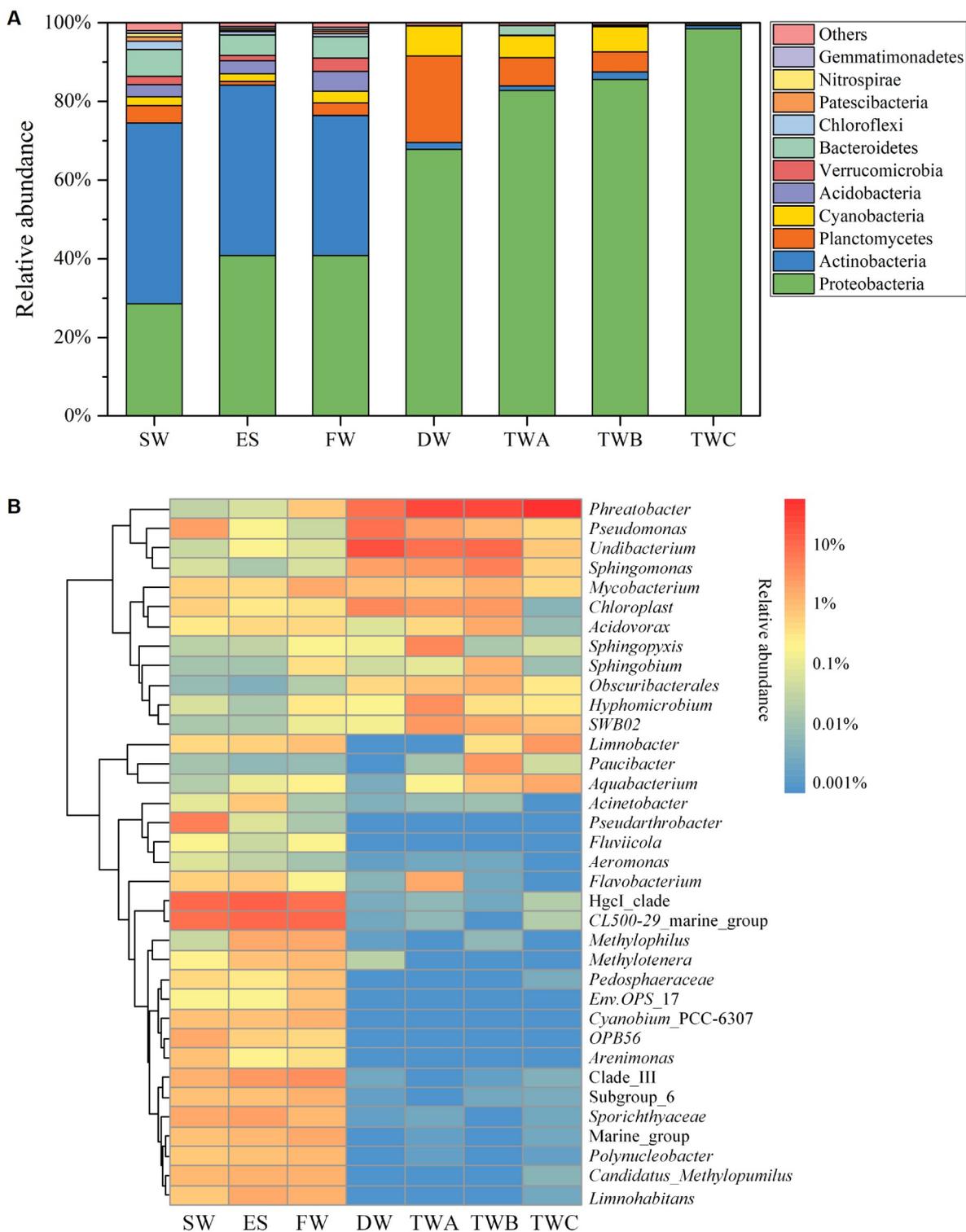


Figure 1 Bacterial community shift at phylum (A) and genus (B) levels in source water (SW), effluent of sedimentation tank (ES), filtered water (FW), chlorine-disinfected water (DW), tap water A (TWA), tap water B (TWB) and tap water C (TWC).

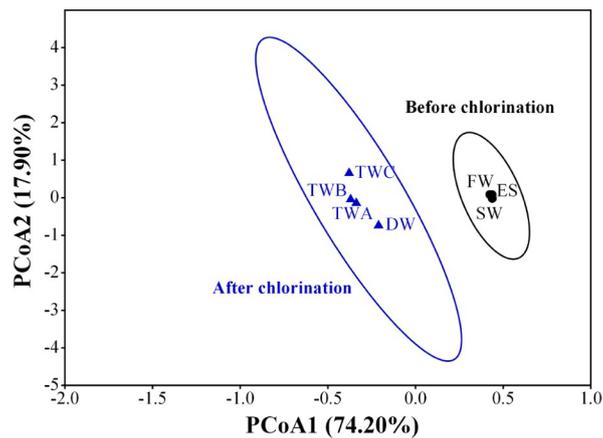


Figure 2 PCoA plot comparing the bacterial communities in source water (SW), effluent of sedimentation tank (ES), filtered water (FW), chlorine-disinfected water (DW), tap water A (TWA), tap water B (TWB) and tap water C (TWC) at genus level.

bacteria phylum, such as *Sphingomonas*, *Phreatobacter*, *Undibacterium*, and *Pseudomonas*, evidently increased after chlorination (Figure 1B).

Persistent and discriminative bacterial genera in drinking water

In this study, 17 bacterial genera concurrently occurred in all the drinking water samples and their relative abundance was over 0.001% (Figure 1B). We defined them as the persistent bacterial genera since they were eliminated by drinking water treatment and existed in the tap water. In general, the persistent bacterial genera occupied $61.09\% \pm 0.04\%$ of the total relative abundance of the detectable bacterial genera in FW, and accounted for $99.93\% \pm 0.00\%$, $97.12\% \pm 0.05\%$, $94.25\% \pm 0.10\%$ and $95.80\% \pm 0.06\%$ of the total relative abundance in the DW, TWA, TWB, and TWC, respectively (Figure 1B). It is worth noting that the total relative abundance of the persistent genera obviously increased after chlorination. Moreover, LEfSe analysis revealed the significant differences between the samples before and after chlorination at several taxonomic levels. In general, a total of 40 bacterial genera showed logarithmic LDA scores higher than 2.0, displaying statistically significant difference after chlorination. For the samples before chlorination, most genera within the phyla Bacteroidetes, Actinobacteria and Acidobacteria were abundant. Several genera belonging to the phylum Planctomycetes were prevalent in the samples after chlorination. The genera *Phreatobacter* and *Sphingomonas* also dominated in the samples after chlorination (Supplementary Figure S5).

Potential pathogenic bacteria in drinking water treatment and distribution systems

The variation of potential pathogenic bacteria in drinking water treatment and distribution systems was observed at both genus (Figure 3A) and species (Figure 3B) levels. Alignment against the human pathogenic bacteria database revealed that 17 genera of potential pathogenic bacteria, including *Pseudomonas*, *Mycobacterium*, *Acinetobacter* and *Aeromonas*, were present in all the water samples (Figure 3A). *Stenotrophomonas*, *Staphylococcus*, *Morganella*, *Escherichia*, *Citrobacter* and *Enterococcus* only existed in SW and then totally disappeared after sedimentation process (Figure 3A). *Acinetobacter* (from $0.010\% \pm 0.018\%$

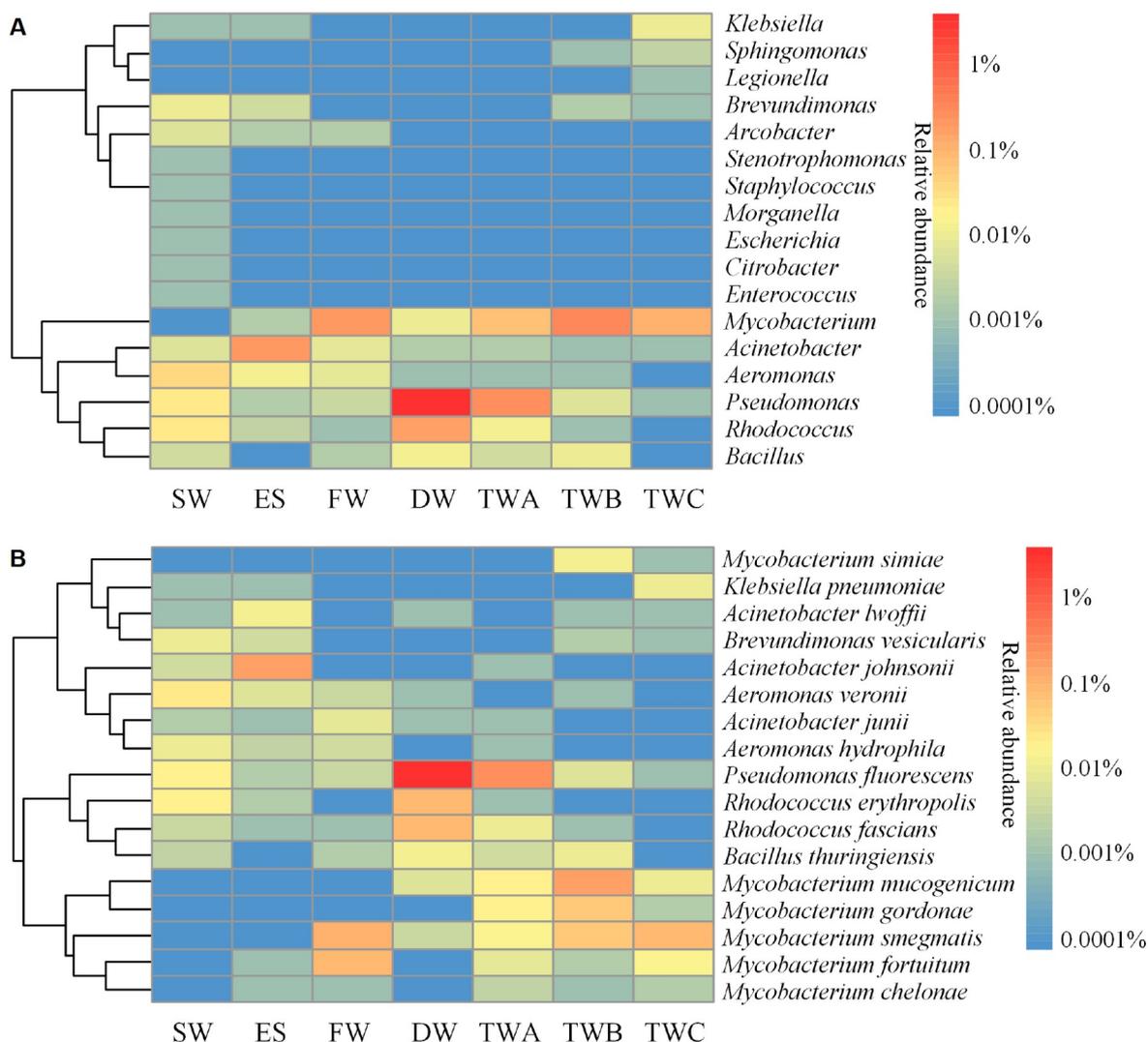


Figure 3 Relative abundance of potential pathogenic bacteria at genus (A) and species (B) (>0.01%) levels in source water (SW), effluent of sedimentation tank (ES), filtered water (FW), chlorine-disinfected water (DW), tap water A (TWA), tap water B (TWB) and tap water C (TWC).

to 0.002%±0.004%), *Aeromonas* (from 0.011%±0.004% to 0.001%±0.002%) and *Mycobacterium* (from 0.252%±0.436% to 0.014%±0.009%) were more sensitive to chlorine since their relative abundance obviously decreased after chlorination (Figure 3A). On the contrary, the relative abundance of *Rhodococcus*, *Pseudomonas*, and *Bacillus* greatly increased after chlorination (Figure 3A). It is worth noting that *Mycobacterium* could gain regrowth after chlorine stress decreased in water distribution pipeline, while *Acinetobacter* and *Aeromonas* could not (Figure 3A). Furthermore, *Mycobacterium*, *Pseudomonas*, *Rhodococcus* and *Bacillus* were also detected in TW samples after pipeline distribution, but *Rhodococcus* and *Bacillus* disappeared in TWC (Figure 3A).

Totally, 56 species of potential pathogenic bacteria were detected in all the water samples, among which 17 had relative abundance of over 0.01% (Figure 3B). Potential pathogenic bacteria *Aeromonas veronii*

dominated in SW, followed by *Rhodococcus erythropolis*, *Pseudomonas fluorescens* and *Aeromonas hydrophila* (Figure 3B). The relative abundance of these bacteria decreased after the processes of sedimentation and filtration (Figure 3B). What is also noteworthy is that chlorination increased the relative abundance of some species of bacteria including *Pseudomonas fluorescens*, *Rhodococcus erythropolis*, *Rhodococcus fascians*, and *Bacillus thuringiensis*, among which *Pseudomonas fluorescens* increased most significantly ($P < 0.05$) (Figure 3B). The most abundant species in TWA, TWB and TWC were *Pseudomonas fluorescens*, *Mycobacterium mucogenicum* and *Mycobacterium smegmatis*, respectively.

Furthermore, 26 species of potential pathogenic bacteria were present in each of TWA, TWB and TWC. The number of potential pathogenic bacteria detected in TWA, TWB and TWC was 13, 19 and 12, respectively (Figure 4). Six potential pathogenic bacteria including *Pseudomonas fluorescens*, *Mycobacterium gordonae*, *Mycobacterium mucogenicum*, *Mycobacterium smegmatis*, *Mycobacterium fortuitum* and *Mycobacterium chelonae* were shared by all the tap water samples, and we defined them as persistent potential pathogenic bacteria, which were closely related to human health and need to be taken seriously. Among the persistent ones, *Pseudomonas fluorescens*, *Mycobacterium mucogenicum* and *Mycobacterium smegmatis* were the most abundant species in TWA, TWB and TWC, respectively (Figure 3B).

Functional profile of bacteria in drinking water

Functional contributions of bacteria in the treatment and distribution systems were explored using PICRUSt software on Galaxy based on OTUs profiles. Results of PICRUSt showed 328 groups at level 3 KEGG Orthology in total, belonging to 41 and 7 groups at level 2 and level 1 categories, respectively. The total

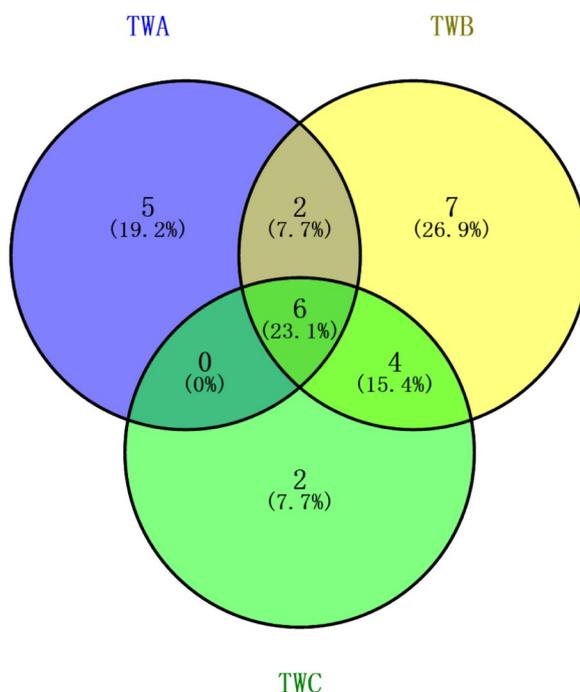


Figure 4 Venn diagram showing the shared and unique potential pathogenic bacteria in tap water A (TWA), tap water B (TWB) and tap water C (TWC).

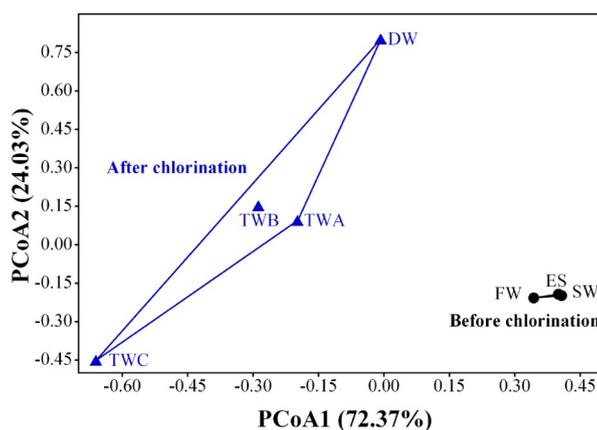


Figure 5 PCoA plot comparing the functional diversity in source water (SW), effluent of sedimentation tank (ES), filtered water (FW), chlorine-disinfected water (DW), tap water A (TWA), tap water B (TWB) and tap water C (TWC).

variation tendency of functional profiles of the bacterial communities before and after chlorination were explored by PCoA (Figure 5), showing no obvious differences among the samples before chlorination. However, great functional changes took place after chlorination and continued in the transportation process. In general, transporters, ABC transporters, DNA repair and recombination proteins, purine metabolism, secretion system, bacterial motility proteins, ribosome, oxidative phosphorylation, and peptidases were the most abundant functional classes at level 3 KEGG Orthology. However, we mainly focused on functional classes involved in human diseases due to their importance to human health (Supplementary Figure S6). Notably, the bacterial communities in drinking water were mainly associated with Alzheimer's disease, Huntington's disease, Tuberculosis, Parkinson's disease, and even pathways in cancer (Supplementary Figure S6). Moreover, results of PICRUSt also revealed that chlorinated water samples, especially after distribution process, harbored the highest amount of human disease functional classes (Supplementary Figure S6).

DISCUSSION

In this study, metagenomic analyses were conducted to find out how the treatment and distribution process influence the diversity and relative abundance of bacterial community in urban drinking water systems. Generally, this study showed that the diversity and abundance of the bacterial populations changed along the treatment and distribution processes, which was confirmed by previous studies demonstrating that the bacterial diversity may be influenced by chlorination disinfection and pipeline distribution [37,38]. Calculation of Chao 1 index, Shannon index and the number of OTUs also revealed the variations of bacterial diversity during the treatment and transportation process. The bacterial diversity was obviously reduced after chlorination and had small fluctuations during the transportation process, and this finding was also reported in a previous study [39]. Notably, chlorination disinfection posed the most considerable influence on the microbial community structure of drinking water. Similarly, previous studies have also indicated that Proteobacteria is the most abundant phylum in water treatment and distribution systems [8,19,40,41]. At the

level of genus, *Phreatobacter* was found to be the most abundant genus in the water samples after chlorination, which agrees with Li *et al.* [15], revealing that *Phreatobacter* spp. dominated in the samples after chlorination. These results indicated that chlorine was the major factor affecting the dynamic composition structure and diversity of drinking water bacterial community. The difference in the sensitivity of bacterial species to chlorine may mainly contribute to the bacterial community shift [42]. Many bacteria within phylum Actinobacteria, such as *Pseudarthrobacter*, were sensitive to chlorine, while *Phreatobacter* in phylum Proteobacteria was not sensitive.

The provision of safe drinking water is necessary for public health. In the past years, opportunistic pathogens have caused numerous outbreaks of waterborne diseases in the developed countries [9]. This study revealed that a total of 17 genera of potential pathogenic bacteria including *Pseudomonas*, *Mycobacterium*, *Acinetobacter*, and *Aeromonas* were detected in the water samples, among which *Pseudomonas* was the most abundant. Similarly, previous studies also reported that *Pseudomonas* had the highest abundance in filtered water in a drinking water treatment plant [4] and the genus also dominated in post-chlorination water samples [39]. *Acinetobacter*, *Aeromonas*, and *Mycobacterium* were supposed to be more sensitive to chlorine since their relative abundances obviously decreased after chlorination. What is also noteworthy is that the relative abundance of *Pseudomonas*, *Rhodococcus*, and *Bacillus* increased greatly after chlorination, indicating that chlorination cannot effectively remove the three genera of potential pathogenic bacteria in drinking water. Likewise, Huang *et al.* [4] showed that chlorination was not effective in removal of *Pseudomonas* in filtered water. Previous studies also demonstrated that the species *Pseudomonas aeruginosa* was frequently detected in post-chlorination drinking water [39,43]. However, in this study, *Pseudomonas fluorescens* was the most abundant bacteria in the post-chlorination samples and was one of the persistent potential pathogenic bacteria in tap water samples. The reason may be that they belong to the same genus and show the similar resistance to chlorine.

As for the persistent potential pathogenic bacteria in tap water samples, several species of *Mycobacteria* were detected, which is similar to the result of a previous study investigating opportunistic pathogens in secondary water supply systems [44]. However, the frequently detectable species *Mycobacterium avium* in drinking water of the United States was not detectable in the water samples in this study [45–48]. This disparity may be related to geographical isolation and the unique local epidemiological characteristics [44]. Moreover, it is interesting that the mycobacteria were not detected in source water but were frequently detected in all the tap water samples, which is similar to the results obtained in a previous study that mycobacteria were commonly detected in treated water [49]. The main reason may include their growth during the process of treatment and distribution, a rather smaller sampling volume of source water than treated water due to clogging of membrane filters, and the random collection of the samples containing no specific mycobacteria due to the heterogeneous distribution of bacterial cells in source water [50]. *Mycobacterium* can cause wound, skin and soft tissue infections, as well as lung disease, especially among children, the elderly, and those with weakened immune systems [51,52]. Thus, it is necessary to further scrutinize mycobacterial species, especially the persistent pathogenic species such as *Mycobacterium chelonae* [46], so as to clarify the exposure risk of associated mycobacteria in local drinking water systems.

Furthermore, functional profiling of the drinking water microbiome using PICRUSt revealed that the bacterial community was involved in various pathways. Although most of the bacteria were related to metabolism and environmental information processing, considerable groups were related to human diseases,

including infectious diseases, metabolic diseases, neurodegenerative diseases, immune system diseases, and even cancers, among which the group of infectious diseases was predominant. PICRUSt-predicted functional profiles inferred relatively little functional variation among the samples before chlorination, suggesting sedimentation and filtration had no great effect on bacterial community function, which may be related to the inconspicuous change of microbial composition and the functional redundancy of drinking water microbial communities [53]. However, the functional profiles changed significantly after chlorination disinfection, due to the variation of bacterial community structure. Although non-culture method like 16S rRNA gene sequencing cannot determine whether the genes associated with human diseases were carried by viable bacteria, there was still a potential risk of infection because of the uptake of free DNA in the drinking water supply piping environment and horizontal gene transfer among bacteria populations [54].

CONCLUSIONS

In conclusion, this study comprehensively revealed a wide and diverse bacterial population and potential pathogenic bacteria in the water samples from urban drinking water treatment and distribution systems, indicating that chlorination was the main factor affecting the bacterial community and potential pathogenic bacteria in drinking water. Chlorination obviously reduced the bacterial diversity, while evidently increased the total relative abundance of the persistent genera. In total, 56 species within 17 genera of potential pathogenic bacteria were found to be present in the drinking water, among which *Pseudomonas fluorescens* and five mycobacteria species were found to be the persistent potential pathogenic bacteria in tap water samples. It is worth noting that a considerable number of the functional profiles of the bacterial community predicted by PICRUSt were related to human diseases, including tuberculosis, Parkinson's disease, and even cancer. Therefore, it is necessary to seek effective methods to control potential pathogenic bacteria, especially *Pseudomonas* and *Mycobacterium*, in drinking water. This study may provide a reference for the future research to reduce the health risks caused by pathogenic bacteria in drinking water.

Data availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Acknowledgements

We thank Dr. Depeng Wang from Nanjing University for providing support for the data analysis.

Funding

This study was financially supported by the National Natural Science Foundation of China (52025102 and 51908274), Nanjing University Excellent Research Program (ZYJH005), and State Key Laboratory of Pollution Control and Resource Reuse Foundation (PCRRF20004).

Author contributions

Ruiming Jiang: conceptualization, methodology, formal analysis, writing-original draft, writing - review & editing, visualization. Zihan Li: software, formal analysis, investigation, writing - review & editing. Qisheng Li: investigation, resources, writing - review & editing. Yunqi Liu: software, formal analysis. Yuning Zhu: methodology, investigation. Zimo

Chen: methodology, investigation. Peng Liu: investigation, resources. Shuyu Jia: conceptualization, writing - review & editing, funding acquisition. Hongqiang Ren: conceptualization, project administration. Xu-Xiang Zhang: conceptualization, data curation, writing - review & editing, supervision, project administration, funding acquisition.

Conflict of interest

The authors declare that there are no conflicts of interest to disclose.

Supplementary information

The supporting information is available online at <https://doi.org/10.1360/nso/20220015>. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

- 1 Huang J, Chen S, Ma X, *et al.* Opportunistic pathogens and their health risk in four full-scale drinking water treatment and distribution systems. *Ecol Eng* 2021; **160**: 106134.
- 2 Jing Z, Lu Z, Mao T, *et al.* Microbial composition and diversity of drinking water: A full scale spatial-temporal investigation of a city in northern China. *Sci Total Environ* 2021; **776**: 145986.
- 3 Sevillano M, Dai Z, Calus S, *et al.* Differential prevalence and host-association of antimicrobial resistance traits in disinfected and non-disinfected drinking water systems. *Sci Total Environ* 2020; **749**: 141451.
- 4 Huang K, Zhang XX, Shi P, *et al.* A comprehensive insight into bacterial virulence in drinking water using 454 pyrosequencing and Illumina high-throughput sequencing. *Ecotoxicol Environ Saf* 2014; **109**: 15–21.
- 5 Hwang C, Ling F, Andersen GL, *et al.* Microbial community dynamics of an urban drinking water distribution system subjected to phases of chloramination and chlorination treatments. *Appl Environ Microbiol* 2012; **78**: 7856–7865.
- 6 Mi Z, Dai Y, Xie S, *et al.* Impact of disinfection on drinking water biofilm bacterial community. *J Environ Sci* 2015; **37**: 200–205.
- 7 Moradi S, Liu S, Chow CWK, *et al.* Developing a chloramine decay index to understand nitrification: A case study of two chloraminated drinking water distribution systems. *J Environ Sci* 2017; **57**: 170–179.
- 8 Liu R, Yu Z, Zhang H, *et al.* Diversity of bacteria and mycobacteria in biofilms of two urban drinking water distribution systems. *Can J Microbiol* 2012; **58**: 261–270.
- 9 Wang H, Edwards MA, Falkinham Joseph O. I, *et al.* Probiotic approach to pathogen control in premise plumbing systems? A review. *Environ Sci Technol* 2013; **47**: 10117–10128.
- 10 Delafont V, Mougari F, Cambau E, *et al.* First evidence of amoebae-mycobacteria association in drinking water network. *Environ Sci Technol* 2014; **48**: 11872–11882.
- 11 Thomas JM, Thomas T, Stuetz RM, *et al.* Your garden hose: A potential health risk due to *Legionella* spp. growth facilitated by free-living amoebae. *Environ Sci Technol* 2014; **48**: 10456–10464.
- 12 Han Z, An W, Yang M, *et al.* Assessing the impact of source water on tap water bacterial communities in 46 drinking water supply systems in China. *Water Res* 2020; **172**: 115469.
- 13 Wang H, Hu C, Zhang S, *et al.* Effects of O₃/Cl₂ disinfection on corrosion and opportunistic pathogens growth in drinking water distribution systems. *J Environ Sci* 2018; **73**: 38–46.
- 14 Thomas JM, Ashbolt NJ. Do free-living amoebae in treated drinking water systems present an emerging health risk? *Environ Sci Technol* 2011; **45**: 860–869.
- 15 Li W, Tan Q, Zhou W, *et al.* Impact of substrate material and chlorine/chloramine on the composition and function of a young biofilm microbial community as revealed by high-throughput 16S rRNA sequencing. *Chemosphere* 2020; **242**: 125310.

- 16 Rajapaksha P, Elbourne A, Gangadoo S, *et al.* A review of methods for the detection of pathogenic microorganisms. *Analyst* 2019; **144**: 396–411.
- 17 Wingender J, Flemming HC. Contamination potential of drinking water distribution network biofilms. *Water Sci Tech* 2004; **49**: 277–286.
- 18 Eichler S, Christen R, Höltje C, *et al.* Composition and dynamics of bacterial communities of a drinking water supply system as assessed by RNA- and DNA-based 16S rRNA gene fingerprinting. *Appl Environ Microbiol* 2006; **72**: 1858–1872.
- 19 Pinto AJ, Xi C, Raskin L. Bacterial community structure in the drinking water microbiome is governed by filtration processes. *Environ Sci Technol* 2012; **46**: 8851–8859.
- 20 Hong PY, Hwang C, Ling F, *et al.* Pyrosequencing analysis of bacterial biofilm communities in water meters of a drinking water distribution system. *Appl Environ Microbiol* 2010; **76**: 5631–5635.
- 21 Pinto AJ, Schroeder J, Lunn M, *et al.* Spatial-temporal survey and occupancy-abundance modeling to predict bacterial community dynamics in the drinking water microbiome. *Mbio* 2014; **5**: e01135–14.
- 22 Proctor CR, Hammes F. Drinking water microbiology—From measurement to management. *Curr Opin Biotechnol* 2015; **33**: 87–94.
- 23 Fang H, Wang H, Cai L, *et al.* Prevalence of antibiotic resistance genes and bacterial pathogens in long-term manured greenhouse soils as revealed by metagenomic survey. *Environ Sci Technol* 2015; **49**: 1095–1104.
- 24 Abia ALK, Alisoltani A, Keshri J, *et al.* Metagenomic analysis of the bacterial communities and their functional profiles in water and sediments of the Apies River, South Africa, as a function of land use. *Sci Total Environ* 2018; **616-617**: 326–334.
- 25 Zhang T, Shao MF, Ye L. 454 Pyrosequencing reveals bacterial diversity of activated sludge from 14 sewage treatment plants. *ISME J* 2012; **6**: 1137–1147.
- 26 Saleem F, Mustafa A, Kori JA, *et al.* Metagenomic characterization of bacterial communities in drinking water supply system of a mega city. *Microb Ecol* 2018; **76**: 899–910.
- 27 Claesson MJ, O’Sullivan O, Wang Q, *et al.* Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. *PLoS ONE* 2009; **4**: e6669.
- 28 Wang Q, Garrity GM, Tiedje JM, *et al.* Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 2007; **73**: 5261–5267.
- 29 Chen L, Xiong Z, Sun L, *et al.* VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. *Nucleic Acids Res* 2012; **40**: D641–D645.
- 30 Qu A, Brulc JM, Wilson MK, *et al.* Comparative metagenomics reveals host specific metavirulomes and horizontal gene transfer elements in the chicken cecum microbiome. *PLoS ONE* 2008; **3**: e2945.
- 31 Oakley BB, Morales CA, Line J, *et al.* The poultry-associated microbiome: Network analysis and farm-to-fork characterizations. *PLoS ONE* 2013; **8**: e57190.
- 32 Huang K, Mao Y, Zhao F, *et al.* Free-living bacteria and potential bacterial pathogens in sewage treatment plants. *Appl Microbiol Biotechnol* 2018; **102**: 2455–2464.
- 33 Langille MGI, Zaneveld J, Caporaso JG, *et al.* Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol* 2013; **31**: 814–821.
- 34 Schloss PD, Westcott SL, Ryabin T, *et al.* Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 2009; **75**: 7537–7541.
- 35 Segata N, Izard J, Waldron L, *et al.* Metagenomic biomarker discovery and explanation. *Genome Biol* 2011; **12**: R60.
- 36 OLIVEROS J. Venny. An interactive tool for comparing lists with Venn’s diagrams. <https://bioinfo.gp.cnb.csic.es/tools/venny/index.html>. 2015.
- 37 Hoffman R, Marshall MM, Gibson MC, *et al.* Prioritizing pathogens for potential future regulation in drinking water. *Environ Sci Technol* 2009; **43**: 5165–5170.
- 38 Mathieu L, Bouteleux C, Fass S, *et al.* Reversible shift in the α -, β - and γ -proteobacteria populations of drinking water

- biofilms during discontinuous chlorination. *Water Res* 2009; **43**: 3375–3386.
- 39 Jia S, Shi P, Hu Q, *et al.* Bacterial community shift drives antibiotic resistance promotion during drinking water chlorination. *Environ Sci Technol* 2015; **49**: 12271–12279.
- 40 Gomez-Alvarez V, Revetta RP, Santo Domingo JW. Metagenomic analyses of drinking water receiving different disinfection treatments. *Appl Environ Microbiol* 2012; **78**: 6095–6102.
- 41 Shi P, Jia S, Zhang XX, *et al.* Metagenomic insights into chlorination effects on microbial antibiotic resistance in drinking water. *Water Res* 2013; **47**: 111–120.
- 42 Chiao TH, Clancy TM, Pinto A, *et al.* Differential resistance of drinking water bacterial populations to monochloramine disinfection. *Environ Sci Technol* 2014; **48**: 4038–4047.
- 43 Wingender J, Flemming HC. Biofilms in drinking water and their role as reservoir for pathogens. *Int J Hyg Environ Health* 2011; **214**: 417–423.
- 44 Li H, Li S, Tang W, *et al.* Influence of secondary water supply systems on microbial community structure and opportunistic pathogen gene markers. *Water Res* 2018; **136**: 160–168.
- 45 Falkinham JO, Norton CD, LeChevallier MW. Factors influencing numbers of *Mycobacterium avium*, *Mycobacterium intracellulare*, and other mycobacteria in drinking water distribution systems. *Appl Environ Microbiol* 2001; **67**: 1225–1231.
- 46 Falkinham III JO. Surrounded by mycobacteria: nontuberculous mycobacteria in the human environment. *J Appl Microbiol* 2009; **107**: 356–367.
- 47 Feazel LM, Baumgartner LK, Peterson KL, *et al.* Opportunistic pathogens enriched in showerhead biofilms. *Proc Natl Acad Sci USA* 2009; **106**: 16393–16399.
- 48 Pfaller SL, Aronson TW, Holtzman AE, *et al.* Amplified fragment length polymorphism analysis of *Mycobacterium avium* complex isolates recovered from southern California. *J Med Microbiol* 2007; **56**: 1152–1160.
- 49 King DN, Donohue MJ, Vesper SJ, *et al.* Microbial pathogens in source and treated waters from drinking water treatment plants in the United States and implications for human health. *Sci Total Environ* 2016; **562**: 987–995.
- 50 Anton V, Rouge P, Daffe M. Identification of the sugars involved in mycobacterial cell aggregation. *Fems Microbiol Lett* 1996; **144**: 167–170.
- 51 Adékambi T. *Mycobacterium mucogenicum* group infections: a review. *Clin Microbiol Infect* 2009; **15**: 911–918.
- 52 Ho YS, Adroub SA, Aleisa F, *et al.* Complete genome sequence of *Mycobacterium fortuitum* subsp *fortuitum* type strain DSM46621. *J Bacteriol* 2012; **194**: 6337–6338.
- 53 Briones A, Raskin L. Diversity and dynamics of microbial communities in engineered environments and their implications for process stability. *Curr Opin Biotechnol* 2003; **14**: 270–276.
- 54 Jutkina J, Rutgersson C, Flach CF, *et al.* An assay for determining minimal concentrations of antibiotics that drive horizontal transfer of resistance. *Sci Total Environ* 2016; **548–549**: 131–138.