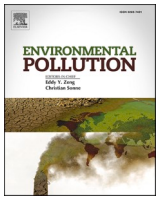
Environmental Pollution 342 (2024) 123066

Contents lists available at ScienceDirect

Environmental Pollution

journal homepage: www.elsevier.com/locate/envpol



Microbial community of municipal drinking water in Hangzhou using metagenomic sequencing☆

Wei Lan [a,b,c], Haiyang Liu [a,b,c], Rui Weng [a,b,c], Yaxiong Zeng [d], Jian Lou [e], Hongxin Xu [e], Yunsong Yu [a,b,c], Yan Jiang [a,b,c,*]

[a] Department of Infectious Diseases, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, 310016, China
[b] Key Laboratory of Microbial Technology and Bioinformatics of Zhejiang Province, Hangzhou, 310016, China
[c] Regional Medical Center for National Institute of Respiratory Diseases, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, 310016, China
[d] College of Chemical and Biological Engineering, Zhejiang University, Hangzhou, 320013, China
[e] Yiwu Water Construction Group Co., Ltd., Yiwu, 322000, China

ARTICLE INFO

Keywords:
Drinking water
Microbial communities
Metagenomic analyses
Functional genes
Virus

ABSTRACT

While traditional culture-dependent methods can effectively detect certain microorganisms, the comprehensive composition of the municipal drinking water (DW) microbiome, including bacteria, archaea, and viruses, remains unknown. Metagenomic sequencing has opened the door to accurately determine and analyze the entire microbial community of DW, providing a comprehensive understanding of DW species diversity, especially in the context of public health concerns during the COVID-19 era. In this study, we found that most of the culturable bacteria and some fecal indicator bacteria, such as *Escherichia coli* and *Pseudomonas aeruginosa*, were non-culturable using culture-dependent methods in all samples. However, metagenomic analysis showed that the predominant bacterial species in the DW samples belonged to the phyla *Proteobacteria* and *Planctomycetes*. Notably, the genus *Methylobacterium* was the most abundant in all water samples, followed by *Sphingomonas*, *Gemmata*, and *Azospirilum*. While low levels of virulence-associated factors, such as the *Esx-5* type VII secretion system (T7SS) and DevR/S, were detected, only the erythromycin resistance gene *erm*(X), an rRNA methyltransferase, was identified at low abundance in one sample. Hosts corresponding to virulence and resistance genes were identified in some samples, including *Mycobacterium* spp. Archaeal DNA (*Euryarchaeota*, *Crenarchaeota*) was found in trace amounts in some DW samples. Viruses such as rotavirus, coxsackievirus, human enterovirus, and SARS-CoV-2 were negative in all DW samples using colloidal gold and real-time reverse transcription polymerase chain reaction (RT−PCR) methods. However, DNA encoding a new order of reverse-transcribing viruses (Ortervirales) and Herpesvirales was found in some DW samples. The metabolic pathways of the entire microbial community involve cell−cell communication and signal secretion, contributing to cooperation between different microbial populations in the water. This study provides insight into the microbial community and metabolic process of DW in Hangzhou, China, utilizing both culture-dependent methods and metagenomic sequencing combined with bioinformatics tools during the COVID-19 pandemic era.

1. Introduction

The importance of access to clean water has been recognized worldwide and is one of the main themes of the UN Sustainable Development Goals. Safe drinking water (DW) is universally acknowledged as a fundamental human right (Mahajna et al., 2022). Clean water is fundamental in the prevention and containment of diseases and the reduced transmission of pathogens and viruses (Bhowmick et al., 2020). It is estimated that more than two billion people face a lack of safely managed DW services (Zvobgo and Do, 2020). Furthermore, some transit processes of treated DW pose potential risks, for example, the uncontrolled proliferation of native bacteria within particles, sediments, and biofilms present in water distribution pipelines (Al-Bedyry et al., 2016).

Active measures have been taken to address the above problems in different countries worldwide. For example, during the late 19th and

☆ This paper has been recommended for acceptance by Sarah Harmon.

* Corresponding author. Department of Infectious Diseases, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, 310016, China.
E-mail address: jiangy@zju.edu.cn (Y. Jiang).

https://doi.org/10.1016/j.envpol.2023.123066
Received 1 September 2023; Received in revised form 1 November 2023; Accepted 27 November 2023
Available online 2 December 2023

early 20th centuries, mortality rates and incidence of waterborne diseases were markedly reduced by implementing water treatment methods such as filtration in many major cities of the US (Gerdes et al., 2023). Currently, many countries, including China, add disinfectants (such as chlorine, mono-chloramine, or chlorine dioxide) to DW as a secondary disinfection step to prevent the regrowth of microbes in finished DW (Priya and Mishra, 2017). However, residuals may introduce selective pressure that may result in the formation of disinfectant-resistant microbial communities (Armstrong et al., 2019). For example, a study on the possible formation of chlorine-tolerant microorganisms in chlorinated water was reported. Certain genera of bacterial populations, including *Sphingomonas*, *Escherichia*, and *Mycobacterium*, in drinking water distribution systems (DWDS) have exhibited varying resistance to monochloramine disinfection (Chiao et al., 2014). Several European countries, including the Netherlands, Germany, Austria, and Switzerland, employ other treatment strategies to enhance DW quality. These strategies focus on removing nutrients that support bacterial growth from the supplied water, thereby limiting the potential for microbial regrowth within the DWDS. The removal of these organic compounds might result in a reduction of carcinogenic byproduct formation. Furthermore, the infrastructure that transports treated DW from the treatment plant to the consumer's tap may cover very long distances, and some facilities are old and contain a motley array of various pipe materials, which might pose other dangerous risks to clean and safe DW (Fish et al., 2017; Prest et al., 2016).

Accurately measuring the complete microbial content of water is a major challenge. Although traditional culture-dependent techniques have played a pivotal role in safeguarding DW, they may not capture the species with extremely low abundances present in the water. A study reported that no more than 1% of the bacterial population can be cultured from oligotrophic systems such as DWDS using conventional cultivation techniques (Falcone-Dias et al., 2015). Metagenomic analysis based on DNA fragments coupled with bioinformatics has allowed the detection, identification, and characterization of all microorganisms in DWDSs in recent years. The inference of the infectious potential of microbial species involves the identification of genes responsible for pathogenic and metabolic attributes (Douterelo et al., 2018). Thus, bacteria, archaea, and viruses can now be detected and identified at the subspecies level.

Here, the microorganisms of five municipal water samples in Hangzhou during the COVID-19 era were determined by traditional culturable methods. Subsequently, we used bioinformatic analysis of the metagenomic sequencing results to recognize the constituents of the whole and special microbial communities of DW. The bacterial components, virulence and resistance-associated factors, archaea, viruses, and metabolic processes of DW were characterized to provide insight into the microbial community and gene functions of DW.

2. Materials and methods

2.1. Sample collection and preparation

Five DW samples were collected, including three municipal tap water samples from distinct families, i.e., JY, LJF, LW, one water sample (SYS) from municipal tap water in the experimental laboratory of a hospital, and one sample (ZJX) from a university. While the exact departure time and date of municipal tap water samples from the water treatment plant (WTP) are not available, the detailed locations and dates of all samples are listed in Fig. S1. All municipal tap water faucets were flushed for 10 min prior to collecting 3000 ml of water within six sterile sampling bags (Hopebio, Qingdao, China). To prevent changes in the microbial community, tap water samples were transported to the laboratory in a cooler box containing ice. The samples were then promptly processed within 1 h of collection. Each sample with 500 ml water was continuously filtered by the fully automatic microbial rapid enrichment system Concentrating Pipette Select™ (InnovaPrep, Drexel, MO, USA) with 0.1 μm PES Flat Filter Concentrating Pipette Tips and 0.05 μm PS Hollow Fiber Filter Concentrating Pipette Tips (InnovaPrep, Drexel, MO, USA) to concentrate bacteria and viruses, respectively. The filter membranes of individual samples were rinsed and preserved at −80 °C until the DNA extraction for metagenomic sequencing. Total residual and free residual chlorines were in line with the national standards of the People's Republic of China "Standards for Drinking Water Quality" (GB5749-2022) (Han et al., 2023).

2.2. Bacterial culture and virus detection

A total of 2500 ml tap water of each sample within five sterile sampling bags was subjected to a heterotrophic plate count, involving filtration onto a 0.45 μm pore size Durapore® polyvinylidene fluoride (PVDF) membranes (Millipore, Billerica, MA, USA), followed by incubation on tryptone soy agar (TSA) or Pseudomonas isolation agar (PA), and identification by mass spectrometry as previously described (Ballesteros et al., 2021). Colony enumeration was conducted at 24-h intervals to ascertain the plate counts at both 30 °C and 37 °C. Rotavirus, coxsackievirus, and human enterovirus were detected by the corresponding immune kit with the colloidal gold technique (Bioneovan Co., Ltd, Beijing, China). Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which causes coronavirus disease 2019 (COVID-19), was detected using real-time fluorescence quantitative polymerase chain reaction (RT−qPCR) assays with kits obtained from Easy Diagnosis Biomedicine Co., Ltd. (Wuhan, China), which was approved by the China Food and Drug Administration (CFDA).

2.3. DNA extraction and metagenomic sequencing

DNA eluted from filter membranes was extracted using a DNeasy PowerSoil Pro Kit (QIAGEN, Hilden, Germany), and its degradation degree and potential contamination were monitored on 1% agarose gels. The purity and concentration of the extracted DNA samples were determined with a Nanodrop 2000 (Thermo Scientific, Waltham, USA) and Qubit 4 Fluorometer (Thermo Scientific, Waltham, USA), respectively. A total of 1 μg DNA per sample was used as input material for the DNA sample preparations. Sequencing libraries were generated using the TIANSeq Fragment/Repair/Tailing module and TIANSeq Fast Ligation module (TIANGEN) following the manufacturer's recommendations, and index codes were added to attribute sequences to each sample. Briefly, the DNA sample was fragmented by the interruption enzyme to a size of 300 bp. DNA fragments were end-polished, A-tailed, and ligated with the full-length adaptor for Illumina sequencing with further PCR amplification. Finally, PCR products were purified (AMPure XP system), and libraries were analyzed for size distribution by Fragment Analyzer and quantified using real-time PCR. After cluster generation, the library preparations were sequenced on an Illumina NovaSeq platform, and 150 bp paired-end reads were generated. To extract high-quality clean reads, quality control measures were adopted, including removing low-quality bases ($Q < 15$ per read 30%) and ambiguous bases ($N > 5\%$ per read), and eliminating the five base pairs that overlapped with adapter sequences. Clean data from each sample were then assembled to produce scaffolds using MEGAHIT v1.1.2 software. Scaffolds were divided into scaftigs without gaps (N). Gene prediction and annotation were performed with scaftigs longer than 500 bp. Metagenomic sequencing generated approximately 7.57×10^7 reads per sample across raw sequence libraries and yielded approximately 7.32×10^7 high-quality clean sequences after removing low-quality reads; detailed information is listed in Table S1.

2.4. Metagenomic sequencing analyses and statistical analysis

Scaftigs (≥500 bp) from each sample were used for open reading frame (ORF) prediction with MetaGeneMark (version 3.26). The predicted ORF (≥100 nt) redundancy was removed using the Cluster

Database at High Identity with Tolerance (CD-HIT (version 4.6)) program to achieve an initial nonredundant gene catalog (nrGC). The reads from the individual samples were realigned to the nrGC using SoapAligner (Version 2.21). The abundance calculation of nonredundant genes (unigenes) involved mapping their read numbers to the nrGC. The relative abundance of unigenes among samples was multiplied by the maximum read number mapping to the nrGC. Exclusive and shared genes were identified by core/pan analysis.

The unigene phylogenetic information was obtained by blasting sequencing reads against the NCBI NR database (including bacteria, fungi, archaea, and viruses, version 17 June 2019) using BLASTP of DIAMOND (version 0.8.1) with E-value $\leq 1 \times 10^{-5}$. Sunburst visualizations of organism-specific k-mer relative abundances (percentages) for each sample were generated using Krona (Ondov et al., 2011). The relative abundance for each taxonomic rank involves summing the relative abundance values of all members. The clustering tree structure of bacteria at different classification levels was based on the Bray–Curtis distance. Unigenes were blasted with the Virulence Factor Database (VFDB) to identify virulence factors with a parameter identity >90% and an E-value $<1 \times e^{-10}$. Unigenes were blasted to the NCBI to identify antimicrobial resistance genes with a parameter identity >90% and coverage >60%. The corresponding hosts were identified by blasting the unigenes filtered with the Pathogen Host Interactions (PHI) database with more than 90% identity. Kyoto Encyclopedia of Genes and Genomes (KEGG) annotation of each gene was performed using BLASTP of DIAMOND against the KEGG database with E-value $\leq 1 \times 10^{-5}$.

3. Results

3.1. Detection of microorganisms and viruses in DW

All water sample membranes were inoculated on TSA culture

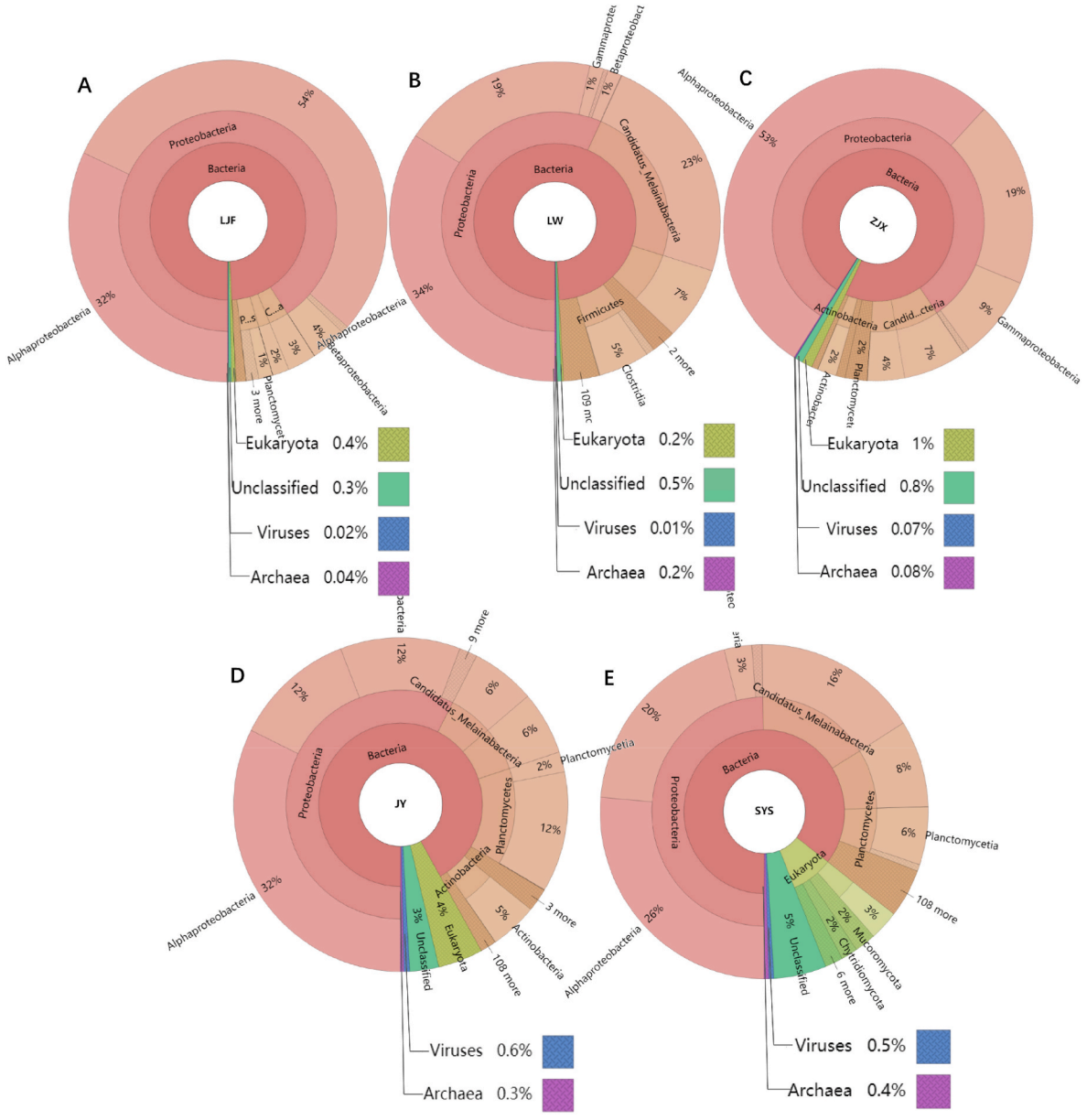


Fig. 1. Krona plot of total species diversity in the municipal tap water. Species composition percentages are shown as the normalized proportion of organism-specific k-mers identified relative to total species diversity detected in the LJF (A), LW (B), ZJX (C), JY (D), and SYS (E) samples. Red, bacteria; teal, eukaryote; purple, archaea; navy blue, virus; green, unclassified. The different classification levels of kidom, phylum, and class associated with organisms are shown in order from the inner cycle to the outer cycle. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

medium, and no known bacteria were collected after 24 h of culture at 30 and 37 °C. None of the DW samples yielded growth, even after incubation for up to 48 h at 30 and 37 °C (Fig. S2). The indicator organisms, *Pseudomonas aeruginosa* and *Escherichia coli*, were not found in any samples by selective culture medium and identification with mass spectrometry (Fig. S3). Viruses, i.e., rotavirus, coxsackievirus, or human enterovirus, were not detectable in any of the water samples with the antigen detection kit employing the colloidal gold method (Fig. S4). RT–qPCR was performed to detect the SARS-CoV-2 genes, and all samples also showed negative results.

3.2. General microbial diversity patterns

To comprehensively and intuitively display the number of genes at different classification levels in each sample, Krona software was used to visualize the species annotation results. The relative abundance of bacteria, archaea, and viruses was visualized using Krona plots. These plots provide a graphical representation of the composition percentages of microbial species in DW samples. All water samples contained a majority of bacteria within the full species diversity (86%–99%), and a minuscule amount of Archaea (0.04–0.4%), eukaryotes (0.2–0.8%), and viruses (0.01–0.6%) (Fig. 1A–E). Except for the SYS sample, the bacterial content of the other four water samples accounted for more than 90% of the total species diversity, specifically in the LW (Figs. 1B and 99%), LJF (Figs. 1A and 99%), ZJX (Figs. 1C and 98%), and JY water samples (Figs. 1D and 92%).

3.3. Bacterial community composition at the phylum level

As shown in Fig. 2, the top 10 microbial species at the phylum level in the five water samples were roughly the same but exhibited distinct distribution patterns. *Proteobacteria* had the highest abundance, accounting for over half of the total composition (49.77–90.82%). *Candidatus_Melainabacteria* accounted for 2.91–23.34%, *Planctomycetes* accounted for approximately 1.28–13.65%, *Firmicutes* accounted for 1.28–13.65%, and *Actinobacteria* accounted for between 0.13 and 4.62%. All these bacteria occupied a prominent position within the top five bacteria at the phylum level. Cluster analysis was conducted to compare the similarity of different samples by constructing a clustering tree based on the Bray–Curtis distance. Here, it was observed that the distribution of microbial relative abundances at the phylum level in the SYS, LJF, and ZJX water samples exhibited similarities, and the JY and LW water samples clustered together.

3.4. Identification of the subclass of the Proteobacteria

Within the phylum *Proteobacteria*, the bulk of *Alphaproteobacteria* was identified in all municipal tap waters, ranging from 53% to 65%, except for the LJF sample (35%). *Betaproteobacteria* accounted for 0.6–4% of all samples (Fig. S5). *Gammaproteobacteria* vary greatly in distinct water samples, ranging from 0.7% to 20%. The abundance of *Deltaproteobacteria* was extremely low in all water samples, ranging from 0.09% to 1% (Fig. S5).

3.5. Bacterial community composition at the genus level

Further classification of microorganisms in all samples was conducted at the class, order, family, genus, and species levels. As shown in Fig. 3, the top 10 microbial species at the genus level in the five water samples were roughly the same but exhibited distinct distribution patterns. *Methylobacterium* was the most abundant (0.23–18.64%) in all water samples, followed by *Sphingomonas* (2.19–14.31%), *Gemmata* (0.06–4.44%), *Azospirilum* (0.03–4.28%), and *Mycobacterium* (0.02–3.92%). *Methylobacterium* and *Sphingomonas* belong to the subclass of *Alpha-Proteobacteria*.

3.6. Virulence and antimicrobial resistance genes

Unigenes were blasted by the VFDB with identity >90% and E-value $<1 \times e^{-10}$, and four virulence genes, including *KasB*, *devR/dosR*, *Esx-5*, and *esxN*, were found in two of five water samples (Table 1). In particular, sample JY contained all of these virulence genes. Blasting the unigenes against the PHI database showed that four virulence genes only corresponded to the *Mycobacterium tuberculosis* (Mtb) and *Mycobacterium marinum* (Mm) hosts. Unigenes were blasted against the antimicrobial resistance gene database of NCBI with the identity >90% and coverage >60%, and only an antibiotic resistance gene (ARG) *'Erm(X)'* encoding a 23S ribosomal RNA methyltransferase with an identity 100% and coverage 100% was found in one of the five water samples. However, after blasting this gene against the PHI database, no host was found due to a low identity of less than 40%.

3.7. Archaea community composition and relative abundance

The proportion of Archaea in the complete microbiome was extremely low (0.04–0.4%). *Euryarchaeota* was the most abundant at the phylum level of the archaea in all DW samples (22–67%), followed by *Candidatus Bathyarchaeota* (1–9%), *Thaumarchaeota* (1–5%), *Candidatus Lokiarchaeota* (0.9–3%), and *Crenarchaeata* (0.9–3%) (Fig. S6, A-E). The archaeal taxonomic subdivisions encompassed 19 classes, 21 orders, 31 families, and 26 genera. Among these, *Methanomicrobia* was the most important class of the *Euryarchaeota* phylum in all samples and had a dominant abundance ranging from 22% to 67%, as detailed in Fig. S7.

3.8. Virus composition and relative abundance

Unigenes were blasted against NCBI, and only two predominant viruses were found. Among them, Ortervirales, a new viral order of reverse-transcribing viruses, occupied 5–34% of the total virus species, and the abundance of Herpesvirales ranged from 11 to 13% in water samples except for the JY sample (Fig. 4, A-E). However, after blasting these two genes against the PHI database, no host was found due to low identity.

3.9. KEGG pathways and functional predictions

Blasting the unigene of each sample against the KEGG pathway

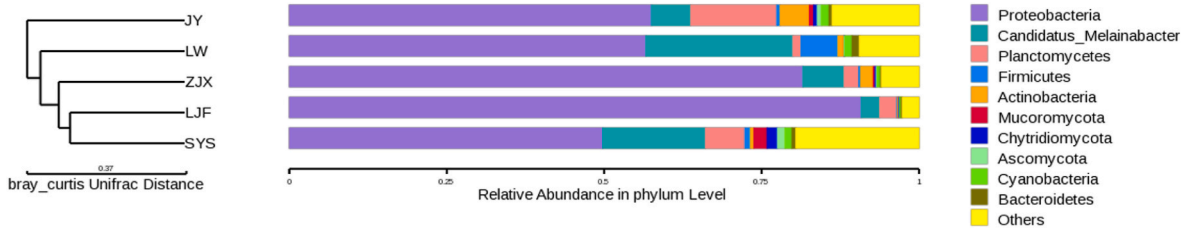


Fig. 2. Clustering tree of bacteria at the phylum level in different DW samples. The left side shows the clustering tree structure based on the Bray–Curtis distance. The right side displays the distribution of the microbial relative abundances at the phylum level for each sample.

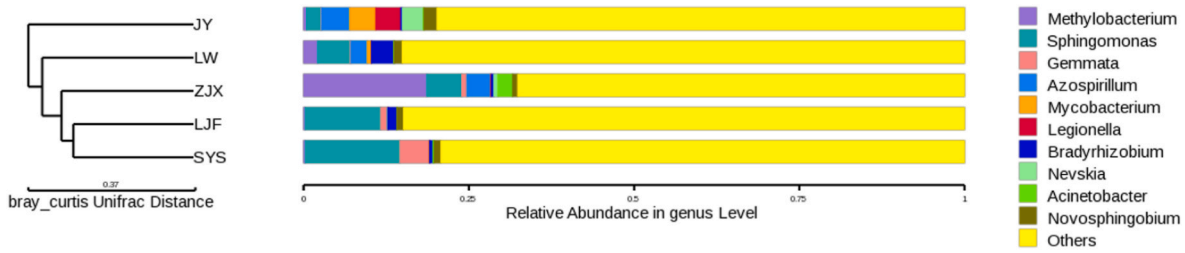


Fig. 3. Clustering tree of bacteria at the genus level in different DW samples. The left side shows the clustering tree structure based on the Bray–Curtis distance. The right side displays the distribution of microbial relative abundances at the genus level for each sample.

Table 1
Virulence factors and corresponding hosts.

Query ID	Identity%	E-value	Virulence factors	Matched host%	Matched Gene name	Matched Pathogen species
JY_200550	95.74	7.00E-48	ESX-5(T7SS)	93.62	EsxO(Rv2346c)	*Mycobacterium_tuberculosis*
JY_238544	95.65	2.00E-33	ESX-5(T7SS)	92.75	EsxO(Rv2346c)	*Mycobacterium_tuberculosis*
LW236468	95.52	2.00E-98	ESX-5(T7SS)	95.52	ESX-5	*Mycobacterium_marinum*
JY_111166	93.00	7.00E-145	ESX-5(T7SS)	93.00	ESX-5	*Mycobacterium_marinum*
JY_428002	92.99	2.00E-110	DevR/S	92.99	dosR_(devR)	*Mycobacterium_tuberculosis*
JY_126064	92.09	5.00E-107	DevR/S	91.63	dosR_(devR)	*Mycobacterium_tuberculosis*
JY_343043	91.46	3.00E-81	DevR/S	91.46	dosR_(devR)	*Mycobacterium_tuberculosis*
JY_35563	90.49	2.00E-164	FAS-II	90.49	kasB	*Mycobacterium_tuberculosis*

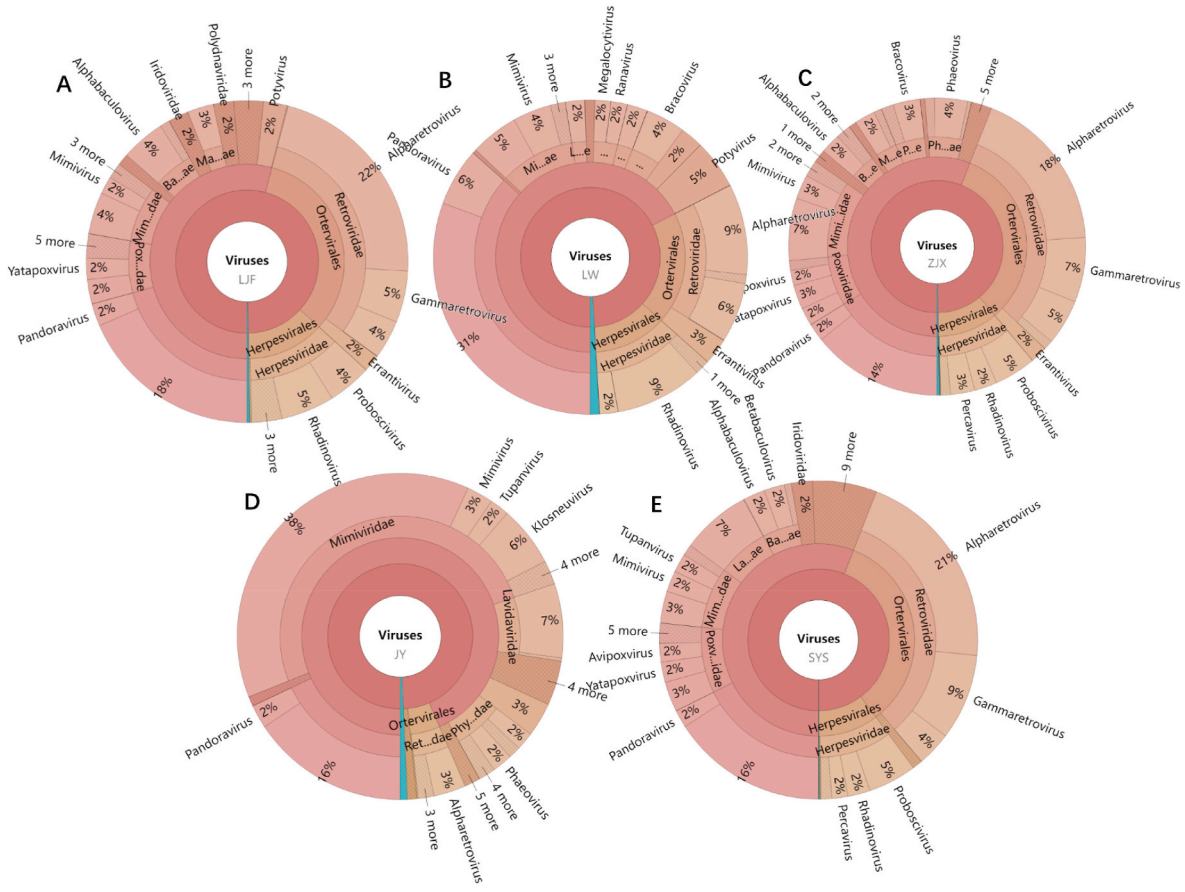


Fig. 4. Krona plot of virus diversity in the municipal tap water samples. The percentages of virus species composition are exhibited as the normalized proportion of organism-specific k-mers observed regarding total virus diversity identified in the LJF (A), LW (B), ZJX (C), JY (D), and SYS (E) samples. The different classification levels of phylum, class, and order related to viruses are shown in order from the inner cycle to the outer cycle.

database, the relative abundance of KEGG pathways at level 1 in all samples almost exhibited the same distribution. BRITE hierarchies and metabolism were the top two relative abundances of the KEGG pathway, accounting for over 65% (Fig. 5, A). At level 2 of the KEGG pathway, 67931, 67242, and 32285 unigenes were substantially enriched in signaling and cellular processing, genetic information processing, and

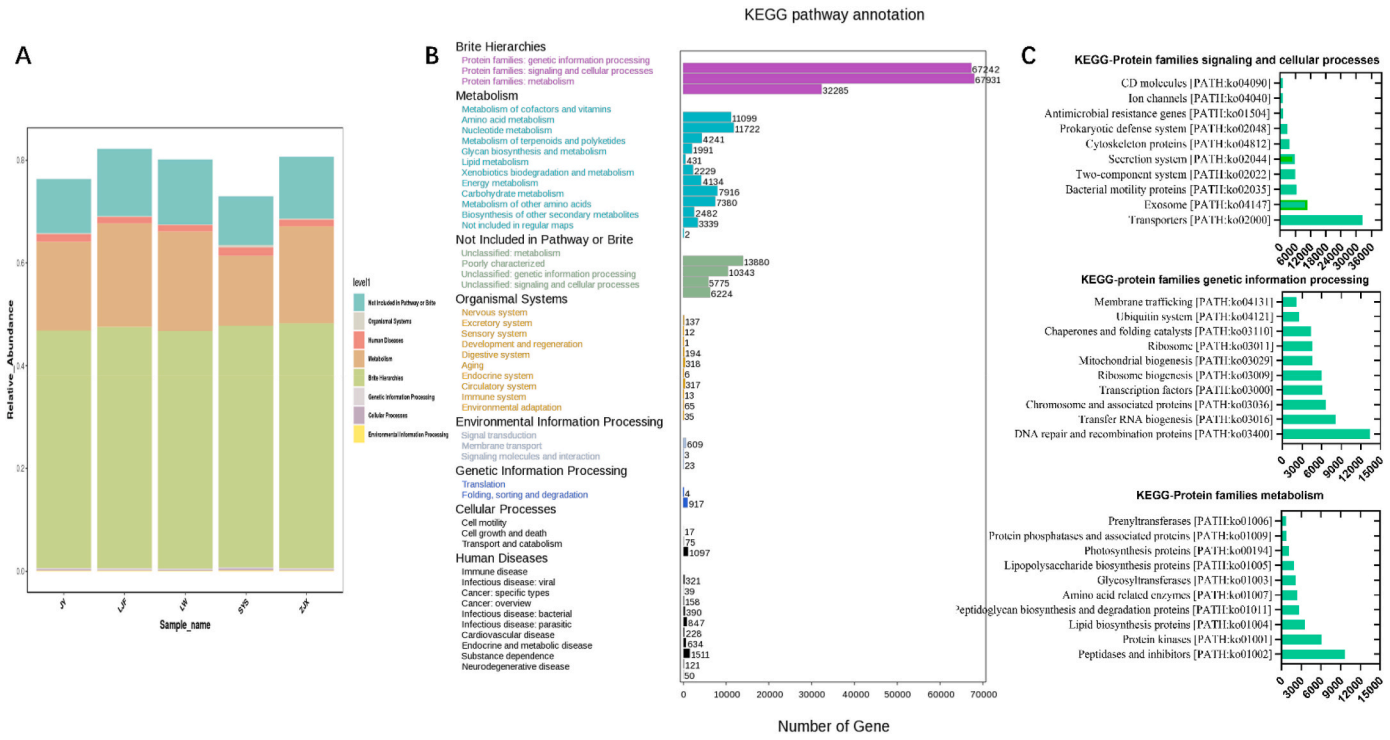


Fig. 5. Gene function enrichment analysis of different classification levels of KEGG pathways in water samples. (A) Relative gene abundance of each water sample at KEGG level 1. (B) Gene annotation at KEGG level 2. (C) Signaling and cellular process, genetic information processing, and metabolism of water samples at KEGG level 3.

metabolism of the BRITE hierarchies. Amino acid metabolism, metabolism of cofactors and vitamins, and energy metabolism were the dominant metabolic processes, including 11722, 11099, and 7916 unigenes, respectively (Fig. 5, B). Furthermore, transporters, exosomes, and bacterial motility proteins were among the top three signaling and cellular processes. DNA repair and recombination proteins, transfer RNA biogenesis, and chromosome and associated proteins occupied the top 3 genetic information processing pathways. Peptidases and inhibitors, protein kinases, and lipid biosynthesis proteins were listed as the top 3 metabolic pathways (Fig. 5, C).

4. Discussion

4.1. Microorganism diversity patterns in DW

Conventionally, it has been believed that culture-dependent methods utilizing indicator microorganisms offer sufficient assurance of water microbial safety. However, these techniques fail to identify all species in the water, particularly non-culturable organisms (Brumfield et al., 2020). In our study, indicator organisms, i.e., *Pseudomonas aeruginosa* and *Escherichia coli*, were non-culturable microorganisms found in all samples. These results were in line with the national standards of the People's Republic of China "Standards for Drinking Water Quality" (GB5749-2022) (Han et al., 2023). Using metagenomic sequencing, non-culturable microorganisms present in the water samples could be identified in our study. Most bacteria and minuscule amounts of archaea and viruses within the total species diversity were identified in all water samples. Several reports of municipal tap water and natural fountain water in the US and France shared similar total microbial compositions with our investigation (Baron et al., 2014; Perrin et al., 2019). The bacterial phyla identified in DW were in agreement with those typically found in DWDSs and natural mineral water (Sala-Comorera et al., 2019), with *Proteobacteria* being the most abundant.

4.2. Bacterial community composition at the phylum level

Utilizing 16S rRNA sequencing, a culture-independent study on the microbiota of a DWDS indicated a prevalence of the subclasses *Alpha-* and *Betaproteobacteria* within the DWDS. Interestingly, the study also revealed that the subclass *Gammaproteobacteria* had a relative abundance of less than 1% (Farhat et al., 2019). Our study also showed that the primary bacterial communities in municipal tap waters were dominated by *Alpha-* and *Betaproteobacteria*, but the abundance of the subclass *Gammaproteobacteria* changed greatly, ranging from 0.7% to 20%, which may be due to the difference between total organic carbon and chlorine residuals in distinct treated water systems. Except for the bacterial phylum *Proteobacteria*, the phyla *Candidatus_Melainabacteria*, *Planctomycetes*, *Firmicutes*, and *Actinobacteria,* which ranged from the top two to five of the bacterial phyla found in all tap water samples, were also examined. A previous study showed that water samples contained higher proportions of *Proteobacteria* and *Bacteroidetes*, while sediment harbored higher proportions of *Planctomycetes*, *Firmicutes*, *Acidobacteria*, *Actinobacteria*, *Chloroflexi*, *Cyanobacteria*, and *Gemmatimonadetes* (Zhang et al., 2021). Here, all samples went through the sediment process of the water treatment system (WTS), which may explain the migration source of *Planctomycetes*, *Firmicutes*, and *Actinobacteria*. Further sequencing research of water and sediment in the WTS needs to be carried out to support this hypothesis.

4.3. Bacterial community composition at the genus level

At the genus level, we revealed that *Methylobacterium* was the most abundant bacterial genus, with proportions ranging from 0.23 to 18.64%, followed by *Sphingomonas* (2.19–14.31%), *Gemmata* (0.06–4.44%), *Azospirilum*, and *Mycobacterium*. *Methylobacterium* and *Sphingomonas* belong to the subclass of *Alpha-Proteobacteria*. The *Methylobacterium* strains may be indigenous environmental bacteria and have been considered to have weak pathogenicity and antibiotic resistance. As demonstrated by some studies, *Methylobacterium*-associated sepsis

has been reported occasionally in AIDS patients and bone marrow-transplant patients (Kovaleva et al., 2014). *Sphingomonas* species are widely distributed in nature and have previously been isolated from soil, wastewater, polluted groundwater, and mineral water (Ma et al., 2017). Here, we found that the appearance of *Sphingomonas* in all water samples might be due to tenacious vitality in the pipeline precipitates or biofilms in municipal DWDSs.

4.4. Virulence factors and hosts

Here, we observed a low prevalence of the bacterial virulence genes *Esx-5* of the type VII secretion system (T7SS) and *DevR/S* in two of five DW samples. The presence of virulence factors in these samples could be linked to the low bacterial content in the water. Furthermore, we found the possible hosts *Mycobacterium tuberculosis* and *Mycobacterium marinum* corresponding to *Esx-5* and *DevR/S* genes by blasting the DNA sequence with the PHI database. The T7SS of *Mycobacteria* encompasses five homologous secretion systems Esx-1 to Esx-5 and plays essential roles in bacterial physiology and host interaction (Bunduc et al., 2021). The Esx-5 system is needed for mycobacterial cell wall stability and host cell lysis. In *Mycobacterium marinum*, the Esx-5 system is needed to induce a caspase-independent form of cell death in macrophages after phagosomal escape of the bacteria, resulting in host immune escape (Shah and Briken, 2016). The DevR-DevS/DosR-DosS two-component system of *Mycobacterium tuberculosis*, which encodes DevS sensor kinase and DevR response regulator, is vital for bacterial adaptation to hypoxia that persists in the environment by inducing dormancy regulon expression. Here, six of the eight hosts belonged to *Mycobacterium tuberculosis*. The appearance of *Mycobacterium tuberculosis* in the water might result from aerosols, which can form spores contributing to their persistence in the environment.

4.5. Antimicrobial resistance genes

Throughout this investigation, only the erythromycin resistance gene *erm*(X), an rRNA methyltransferase, was detectable at an abundance level above the limit of confidence. Some molecular studies of DWDSs have reported the presence of specific pathogenic bacteria that carry ARGs (Gomes et al., 2020; Hu et al., 2021). However, a study in which ARGs were not detectable in DW was also reported (Brumfield et al., 2020). Here, only one ARG in all samples was detectable, suggesting that microfiltration and reverse osmosis are effective membrane-based processes that can remove both whole cells and transmissible genetic elements from water sources. Erm(X) can protect the ribosome from inactivation due to antibiotic binding in bacteria. In our study, the *erm* (X) gene was detectable, but its corresponding host was not matched. This gene may be from the transfer of the other bacterium, as evidenced by a study that showed that genomic islands play an essential role in the dissemination of the gene *erm*(X) among some species, i.e., the *Bifidobacterium* species (Wang et al., 2017).

4.6. Archaea community

The abundance of archaea ranged from 0.04 to 0.4% in this study, which is consistent with other studies that reported that the content of archaea was lower than 1%. The overall higher proportion of the phylum *Euryarchaeota* in all samples is related to the higher water column depth in the pipeline, which may contribute microenvironments to reducing the interaction between the water and sediment, producing more favorable conditions for the success of anaerobic groups (Silveira et al., 2021). The difference in abundance of the phylum *Crenarchaeota* in different samples may be associated with the oxygen content in the transit pipeline (Choi et al., 2022). *Euryarchaeota* is composed of methanogenic members such as *Methanomicrobia*, *Methanosaeta*, *Methanocella*, and other taxa groups with anaerobic metabolism (Berdjeb et al., 2013). Here, *Methanomicrobia* was the dominant phylum of *Euryarchaeota*, and the discrepancy in distinct samples may be caused by water salinity and pH.

4.7. Virus composition

During the COVID-19 period, safe and clean DW could effectively block the water-mediated transmission of viruses, and the detection and analysis of the virus content in DW are very important in protecting public health. Several viruses, e.g., enterovirus, rotavirus, norovirus, adenovirus, and hepatitis A and E viruses, can cause various human infections, including acute gastroenteritis (Gall et al., 2015). Here, rotavirus, coxsackievirus, human enterovirus, and SARS-CoV-2 were not detectable in any of the water samples, and these results were in line with the national standards of the People's Republic of China "Standards for Drinking Water Quality" (GB5749-2022). In addition, we detected DNA fragments of Ortervirales, a new virus order unifying five families of reverse-transcribing viruses, and all these viruses utilize host tRNA molecules as primers for genome replication by reverse transcription (Krupovic et al., 2018). In addition, the DNA encoding Herpesviruses, which are host-specific pathogens that are widespread among vertebrates and can establish long-term latency and manipulation of the host immune response, were identified. However, no corresponding hosts were found in our DW samples, and the function of these viruses needs further study. Most viruses of prokaryotes possess double-stranded (ds) DNA genomes, while RNA viruses are limited in the presence of eukaryotes (Koonin et al., 2015). Here, we did not find the RNA virus, which is due to only the DNA fragments isolated from the samples.

4.8. KEGG pathways and functions

The metabolic processes, which contribute to the survival status of different species in a water environment, were characterized for all species. Here, similar patterns of the KEGG metabolism profile in all samples could indicate many shared functional traits. At a finer annotation resolution or KEGG level 2 or 3, transporters, exosomes, and bacterial motility proteins of the signaling and cellular process, DNA repair and recombination proteins, transfer RNA biogenesis, and chromosome and associated proteins of the genetic information processing, and peptidases and inhibitors, protein kinases, and lipid biosynthesis proteins of the metabolism were identified. These pathways could involve cell−cell communication and signal secretion, contributing to cooperation between microbial populations in the water (Abisado et al., 2018; Petrova et al., 2023). Although it is not yet possible to determine the corresponding metabolic profile of every species due to the limitations of sequencing methods, these results provide an avenue to evaluate the role of the total species.

5. Conclusions

The combination of traditional culture-dependent methods and metagenomic sequencing coupled with bioinformatics tools presents an effective approach for investigating the autochthonous microbial community within DW during the COVID-19 era. The following conclusions could be drawn from the results of this study:

(1) Although culturable bacteria were not detected using traditional methods, the phyla *Proteobacteria* and *Planctomycetes* and the genus *Methylobacterium* were the most predominant bacterial taxa in DW using metagenomics.
(2) *Mycobacterium* spp., hosts corresponding to virulence factor '*Esx-5*', was identified by metagenomics.
(3) While viruses were not found in water isolated from the COVID-19 era, a new order of reverse-transcribing viruses Ortervirales was identified in the water using metagenomics.
(4) The identified metabolic pathways reflect the cooperation between different microbial populations in the water.

The results of this study provide insight into bacteria, viruses, virulence-associated factors, and metabolic processes present in DW from various regions, yielding new information on the microbial community structure of tap water and its associated gene functions. In the post-COVID-19 era, effective surveillance of clean and safe sanitation water is very important to the protection of public health.

CRediT authorship contribution statement

Wei Lan: Conceptualization, Sample collection, Formal analysis, Visualization, Writing - original draft. **Haiyang Liu:** Conceptualization, Sample collection, Bacteria culture, Virus detection. **Rui Weng:** Bacteria culture and Microorganism identification. **Yaxiong Zeng:** Sample collection. **Jian Lou:** Sample collection. **Hongxin Xu:** Sample collection. **Yunsong Yu:** Supervision, Funding acquisition, Project administration. **Yan Jiang:** Supervision, Writing - review & editing, Project administration, Funding acquisition, All authors contributed to the article and approved the submitted version.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by the Key Research Program of the Science Technology Department of Zhejiang Province (Grant No. 2021C03179). We sincerely thank Jingfeng Luo (LJF) and Junxin Zhou (ZJX) of Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, for their sampling and technical help.

Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envpol.2023.123066.

References

Abisado, R.G., Benomar, S., Klaus, J.R., Dandekar, A.A., Chandler, J.R., 2018. Bacterial quorum sensing and microbial community interactions. mBio 9, e02331-17.
Al-Bedyry, N.K., Sathasivan, A., Al-Ithari, A.J., 2016. Ranking pipes in water supply systems based on potential to cause discolored water complaints. Process. Saf. Environ. Prot 104, 517–522.
Armstrong, H., Alipour, M., Valcheva, R., Bording-Jorgensen, M., Jovel, J., Zaidi, D., Shah, P., Lou, Y., Ebeling, C., Mason, A.L., Lafleur, D., Jerasi, J., Wong, G.K.S., Madsen, K., Carroll, M.W., Huynh, H.Q., Dieleman, L.A., Wine, E., 2019. Host immunoglobulin G selectively identifies pathobionts in pediatric inflammatory bowel diseases. Microbiome 7, 1.
Ballesteros, H.G.F., Rosman, A.C., Carvalho, T.L.G., Grativol, C., Hemerly, A.S., 2021. Cell wall formation pathways are differentially regulated in sugarcane contrasting genotypes associated with endophytic diazotrophic bacteria. Planta 254, 109.
Baron, J.L., Vikram, A., Duda, S., Stout, J.E., Bibby, K., 2014. Shift in the microbial ecology of a hospital hot water system following the introduction of an on-site monochloramine disinfection system. PLoS One 9, e102679.
Berdjeb, L., Pollet, T., Chardon, C., Jacquet, S., 2013. Spatio-temporal changes in the structure of archaeal communities in two deep freshwater lakes. FEMS Microbiol. Ecol. 86, 215–230.
Bhowmick, G.D., Dhar, D., Nath, D., Ghangrekar, M.M., Banerjee, R., Das, S., Chatterjee, J., 2020. Coronavirus disease 2019 (COVID-19) outbreak: some serious consequences with urban and rural water cycle. NPJ Clean Water 3, 32.
Brumfield, K.D., Hasan, N.A., Leddy, M.B., Cotruvo, J.A., Rashed, S.M., Colwell, R.R., Huq, A., 2020. A comparative analysis of drinking water employing metagenomics. PLoS One 15, e0231210.
Bunduc, C.M., Fahrenkamp, D., Wald, J., Ummels, R., Bitter, W., Houben, E.N.G., Marlovits, T.C., 2021. Structure and dynamics of a mycobacterial type VII secretion system. Nature 593, 445–448.
Chiao, T.-H., Clancy, T.M., Pinto, A., Xi, C., Raskin, L., 2014. Differential resistance of drinking water bacterial populations to monochloramine disinfection. Environ. Sci. Technol. 48, 4038–4047.
Choi, H., Lee, H., Kim, D.-H., Lee, K.-K., Kim, Y., 2022. Physicochemical and isotopic properties of ambient aerosols and precipitation particles during winter in Seoul, South Korea. Environ. Sci. Pollut. Res. Int. 29, 11990–12008.
Douterelo, I., Calero-Preciado, C., Soria-Carrasco, V., Boxall, J.B., 2018. Whole metagenome sequencing of chlorinated drinking water distribution systems. Environ. Sci.: Water Res. Technol. 4, 2080–2091.
Falcone-Dias, M.F., Centron, D., Pavan, F., Moura, A.C.d.S., Naveca, F.G., De Souza, V.C., Farache Filho, A., Leite, C.Q.F., 2015. Opportunistic pathogens and elements of the resistome that are common in bottled mineral water support the need for continuous surveillance. PLoS One 10, e0121284.
Farhat, M., Alkharsah, K.R., Alkhamis, F.I., Bukharie, H.A., 2019. Metagenomic study on the composition of culturable and non-culturable bacteria in tap water and biofilms at intensive care units. J. Water Health 17, 72–83.
Fish, K., Osborn, A.M., Boxall, J.B., 2017. Biofilm structures (EPS and bacterial communities) in drinking water distribution systems are conditioned by hydraulics and influence discolouration. Sci. Total Environ. 593-594, 571–580.
Gall, A.M., Mariñas, B.J., Lu, Y., Shisler, J.L., 2015. Waterborne viruses: a Barrier to safe drinking water. PLoS Pathog. 11, e1004867.
Gerdes, M.E., Miko, S., Kunz, J.M., Hannapel, E.J., Hlavsa, M.C., Hughes, M.J., Stuckey, M.J., Francois Watkins, L.K., Cope, J.R., Yoder, J.S., Hill, V.R., Collier, S.A., 2023. Estimating waterborne infectious disease Burden by exposure Route, United States, 2014. Emerg. Infect. Dis. 29, 1357–1366.
Gomes, I.B., Maillard, J.-Y., Simões, L.C., Simões, M., 2020. Emerging contaminants affect the microbiome of water systems-strategies for their mitigation. NPJ Clean Water 3, 39.
Han, J., Zhang, L., Ye, B., Gao, S., Yao, X., Shi, X., 2023. The standards for drinking water quality of China (2022 edition) will Take effect. China CDC Weekly 5, 297–300.
Hu, Y., Jiang, L., Sun, X., Wu, J., Ma, L., Zhou, Y., Lin, K., Luo, Y., Cui, C., 2021. Risk assessment of antibiotic resistance genes in the drinking water system. Sci. Total Environ. 800, 149650.
Koonin, E.V., Dolja, V.V., Krupovic, M., 2015. Origins and evolution of viruses of eukaryotes: the ultimate modularity. Virology 479–480, 2–25.
Kovaleva, J., Degener, J.E., van der Mei, H.C., 2014. Methylobacterium and its role in health care-associated infection. J. Clin. Microbiol. 52, 1317–1321.
Krupovic, M., Blomberg, J., Coffin John, M., Dasgupta, I., Fan, H., Geering Andrew, D., Gifford, R., Harrach, B., Hull, R., Johnson, W., Kreuze Jan, F., Lindemann, D., Llorens, C., Lockhart, B., Mayer, J., Muller, E., Olszewski Neil, E., Pappu Hanu, R., Pooggin Mikhail, M., Richert-Pöggeler Katja, R., Sabanadzovic, S., Sanfaçon, H., Schoelz James, E., Seal, S., Stavolone, L., Stoye Jonathan, P., Teycheney, P.-Y., Tristem, M., Koonin Eugene, V., Kuhn Jens, H., 2018. Ortervirales: new virus order unifying five families of reverse-transcribing viruses. J. Virol. 92, e00515-18.
Ma, L., Li, B., Jiang, X.-T., Wang, Y.-L., Xia, Y., Li, A.-D., Zhang, T., 2017. Catalogue of antibiotic resistome and host-tracking in drinking water deciphered by a large scale survey. Microbiome 5, 154.
Mahajna, A., Dinkla, I.J.T., Euverink, G.J.W., Keesman, K.J., Jayawardhana, B., 2022. Clean and safe drinking water systems via metagenomics data and Artificial intelligence: state-of-the-Art and future perspective. Front. Microbiol. 13, 832452.
Ondov, B.D., Bergman, N.H., Phillippy, A.M., 2011. Interactive metagenomic visualization in a Web browser. BMC Bioinf. 12, 385.
Perrin, Y., Bouchon, D., Delafont, V., Moulin, L., Héchard, Y., 2019. Microbiome of drinking water: a full-scale spatio-temporal study to monitor water quality in the Paris distribution system. Water Res. 149, 375–385.
Petrova, O., Parfirova, O., Gogoleva, N., Vorob'ev, V., Gogolev, Y., Gorshkov, V., 2023. The role of intercellular signaling in the regulation of bacterial adaptive proliferation. Int. J. Mol. Sci. 24, 7266.
Prest, E.I., Hammes, F., Kötzsch, S., van Loosdrecht, M.C.M., Vrouwenvelder, J.S., 2016. A systematic approach for the assessment of bacterial growth-controlling factors linked to biological stability of drinking water in distribution systems. Water Supply 16, 865–880.
Priya, T., Mishra, B.K., 2017. Enzyme mediated chloroform biotransformation and quantitative cancer risk analysis of trihalomethanes exposure in South East Asia. Exposure Health 9, 61–75.
Sala-Comorera, L., Blanch, A.R., Casanovas-Massana, A., Monleón-Getino, A., García-Aljaro, C., 2019. Traceability of different brands of bottled mineral water during shelf life, using PCR-DGGE and next generation sequencing techniques. Food Microbiol. 82, 1–10.
Shah, S., Briken, V., 2016. Modular organization of the ESX-5 secretion system in *Mycobacterium tuberculosis*. Front. Cell. Infect. Microbiol. 6, 49.
Silveira, R., Silva, M.R.S.S., de Roure Bandeira de Mello, T., Alvim, E.A.C.C., Marques, N. C.S., Kruger, R.H., da Cunha Bustamante, M.M., 2021. Bacteria and archaea communities in Cerrado natural pond sediments. Microb. Ecol. 81, 563–578.
Wang, N., Hang, X., Zhang, M., Peng, X., Yang, H., 2017. New genetic environments of the macrolide-lincosamide-streptogramin resistance determinant *erm*(X) and their influence on potential horizontal transferability in bifidobacteria. Int. J. Antimicrob. Agents 50, 572–580.
Zhang, L., Delgado-Baquerizo, M., Shi, Y., Liu, X., Yang, Y., Chu, H., 2021. Co-existing water and sediment bacteria are driven by contrasting environmental factors across glacier-fed aquatic systems. Water Res. 198, 117139.
Zvobgo, L., Do, P., 2020. COVID-19 and the call for 'Safe Hands': Challenges facing the under-resourced municipalities that lack potable water access-A case study of Chitungwiza municipality, Zimbabwe. Water Res. 9, 100074.