Disruptive effects of sewage intrusion into drinking water

Microbial succession and organic transformation at molecular level

Fan, Mengqing; Ren, Anran; Yao, Mingchen; Li, Xiaoming; van der Meer, Walter; Yu, Guo; Medema, Gertjan; Rose, Joan; Liu, Gang

DOI
10.1016/j.watres.2024.122281

Publication date
2024

Document Version
Final published version

Published in
Water Research

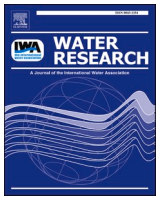
Citation (APA)
Fan, M., Ren, A., Yao, M., Li, X., van der Meer, W., Yu, G., Medema, G., Rose, J., & Liu, G. (2024). Disruptive effects of sewage intrusion into drinking water: Microbial succession and organic transformation at molecular level. *Water Research*, *266*, Article 122281. https://doi.org/10.1016/j.watres.2024.122281

Water Research 266 (2024) 122281

Contents lists available at ScienceDirect

Water Research

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



Disruptive effects of sewage intrusion into drinking water: Microbial succession and organic transformation at molecular level

Mengqing Fan [a,b,c], Anran Ren [a,b,c], Mingchen Yao [a,b,c], Xiaoming Li [a,b], Walter van der Meer [d,e], Guo Yu [f], Gertjan Medema [c,g], Joan Rose [h], Gang Liu [a,b,c,*]

[a] Key Laboratory of Drinking Water Science and Technology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing, 100085, China
[b] University of Chinese Academy of Sciences, Beijing, 100049, China
[c] Sanitary Engineering, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, the Netherlands
[d] Membrane Science and Technology, Faculty of Science and Technology, Twente University, the Netherlands
[e] Oasen Drinkwater, Gouda, the Netherlands
[f] Guangxi Key Laboratory of Environmental Pollution Control Theory and Technology, Guilin University of Technology, Guilin, 541004, China
[g] KWR Watercycle Research Institute, PO Box 1072, 3430 ,BB Nieuwegein, the Netherlands
[h] Department of Fisheries and Wildlife, Michigan State University, East Lansing, MI, 48823, USA

ARTICLE INFO

Keywords:
Sewage intrusion
Drinking water
Bacterial growth
Dissolved organic matter
Disruptive effects

ABSTRACT

Drinking water distribution systems are increasingly vulnerable to sewage intrusion due to aging water infrastructure and intensifying water stress. While the health risks associated with sewage intrusion have been extensively studied, little is known about the impacts of intruded bacteria and dissolved organic matter (DOM) on microbiology in drinking water. In this dynamic study, we demonstrate that the intrusion of 1 % sewage into tap water resulted in immediate contamination, including an 8-fold increase in biomass (TCC), a 48.9 % increase in bacterial species (ASVs), a 12.5 % increase in organic carbon content (DOC), and a 13.5 % increase in unique DOM molecular formulae. Over time, sewage intrusion altered tap water microbiology by accelerating bacterial growth rates (5-fold faster), selectively promoting ASVs in community succession, and producing 998 more unique DOM formulae. More significantly, statistical analysis revealed that the intrusion of 1 % sewage shifted the driving force of bacterial and DOM composition covariance from a DOM-dependent process in tap water to a bacterial-governed process post-intrusion. Our results clearly demonstrate the disruptive effects of sewage intrusion into tap water, emphasizing the urgent need to consider the long-lasting impacts of sewage intrusion in drinking water distribution systems, in addition to its immediate health risks.

1. Introduction

Intrusion in drinking water distribution systems (DWDSs) is defined as a specialized back-flow of non-potable water into DWDSs (LeChevallier et al., 2003), resulting from breaches in physical and hydraulic integrity (Besner et al., 2011; NRC, 2006). DWDSs face heightened vulnerability to intrusion due to frequent compromised physical integrity (ASCE, 2022; Barton et al., 2019) (e.g., circumferential breaks, longitudinal splits, joint failure, and holes), diminished hydraulic integrity (Besner et al., 2010; Hunter et al., 2005; Nygard et al., 2007) (e.g., breaks and maintenance work, intermittent water supply), and transient negative pressure (Fontanazza et al., 2015; Karim et al., 2003) (e.g., intermittent pump activation, opening/closing of hydrants, valve slam/failures). In these scenarios, sewage emerges as the most prevalent intrusion source, given the high volume of sewage discharging directly into environment, and the close proximity of underground sewage and drinking water pipelines (Haydar et al., 2009; Kauppinen et al., 2019; M. Aenab and K. Singh, 2012). Sewage introduces contaminants, such as pathogens and heavy metals, into DWDSs, which could rapidly reach tens of thousands of consumers within hours to days, posing significant threats to both public health and the environment. However, as water infrastructures age (Agency, 2011; Barton et al., 2019) and water stress intensifies (Semenza, 2020), the intrusion of sewage into drinking water and the associated risks are expected to further deteriorate.

The cases of sewage intrusion into drinking water may not be extensively documented individually due to the localized nature of such

* Corresponding author at: Chinese Academy of Science; Delft University of Technology, China.
E-mail address: gliu@rcees.ac.cn (G. Liu).

https://doi.org/10.1016/j.watres.2024.122281
Received 5 June 2024; Received in revised form 28 July 2024; Accepted 15 August 2024
Available online 27 August 2024

incidents and privacy considerations. Nevertheless, on a global scale, there is substantial evidence indicating that sewage-contaminated drinking water has led to outbreaks of waterborne diseases and raised health concerns. To illustrate, in Seoul, Republic of Korea, from May to June 2012, a waterborne outbreak of 124 cases of cryptosporidiosis occurred in an older high-rise apartment complex's plumbing system, associated with sewage-contaminated drinking water (Cho et al., 2013). In May 2015, Kasese District in Uganda witnessed a prolonged cholera outbreak caused by sewage-contaminated drinking water. This outbreak affected over 80 villages, resulting in 183 suspected cases and 2 fatalities (Kwesiga et al., 2017). More recently, instances of sewage intrusion into DWDSs in Finland led to outbreaks involving multiple waterborne pathogenic microbes. In 2016 and 2018, these incidents impacted 790 and 4000 people, respectively (Kauppinen et al., 2019). These cases underscore the importance of addressing sewage contamination in drinking water systems to mitigate the risk of waterborne diseases and safeguard public health.

Amid recurring health concerns, the intrusion of sewage in drinking water systems has become a notable issue. Previous studies focused on the timely detection of intrusion events (Heibati et al., 2017; Prest et al., 2013; Sorensen et al., 2021; Vang et al., 2014) and the prediction of infection risks (Odhiambo et al., 2023; Teunis et al., 2010; Vinas et al., 2022). However, it has been neglected that microbes and dissolved organic matter (DOM) introduced by sewage intrusion would not remain stagnant or isolated from drinking water microbes and organics post-intrusion. Instead, they undergo growth and transformation, potentially giving rise to disruptive impacts on drinking water microbiology. For example, previous studies have discovered co-variations between microbes and organic matter in aquatic environments such as oceans, rivers, and lakes (Osterholz et al., 2016; Yang et al., 2020b; Zhou et al., 2024). Additionally, studies have indicated that such chemical and biological covariance can influence ecosystem functions (Logue et al., 2016; Tanentzap et al., 2019), which is important for the biosafety and biostability of drinking water. Therefore, it is essential to monitor the synergistic changes in microbial succession and organic transformation to better understand the disruptive effects following sewage intrusion. However, crucial knowledge in this regard is lacking, leaving a fundamental gap in our understanding of the aftermath of sewage intrusion in drinking water. Particularly, we lack insights into the disruptive effects induced by microbes and organics in sewage on the growth dynamics and evolutionary mechanisms of drinking water microbes.

For sewage intrusion events detection, both microbes and dissolved organic matter (DOM) are sensitive indicators. It has been reported that adenosine triphosphate (ATP) and flow cytometry total cell counts (TCC) can detect the occurrence of contamination by detecting increases in biomass (Prest et al., 2013; Vang et al., 2014), while fluorescence spectroscopy can detect changes in the fluorescent components of DOM in drinking water (Heibati et al., 2017; Sorensen et al., 2021). By ATP/TCC and DOM based methods, as low as 0.01 % - 2 % intrusion is detectable. However, these quantitative methods can only reveal changes in concentrations of microbes and DOM but not their compositions. The development and application of high-throughput sequencing and Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR MS) offered new possibilities to qualitatively study the covariance of microbes and DOM at molecular level (Antony et al., 2017; Zark and Dittmar, 2018).

The above-mentioned sewage intrusion studies were all conducted in laboratory, because it is virtually impossible to catch field intrusion events or simulate them in field environment due to their high complexity and limited accessibility of drinking water distribution systems. In the field environment, the occurrence of sewage intrusion is either delayed in detection or under spotlight only after associated infections were reported publicly. Injecting sewage into field drinking water pipes would be impossible and even illegal because of its associated public health threats. In this study, to address the critical question of the changes of microbial succession and DOM transformation induced by sewage intrusion at molecular level, we designed a simulation experiment with specific laboratory conditions. For simulating the worst scenario and reveal the dynamic changes of microbes and DOM, sewage intrusion was monitored for a period of seven days. The quantitative changes of microbes and DOM are measured by TCC and dissolved organic carbon (DOC), while the microbial community is assessed by Illumina sequencing, and the DOM is characterized by FT-ICR MS. The combination of these high-throughput and ultrahigh-resolution techniques provided profound insights into the disruptive changes arising from sewage intrusion, including microbial succession and organic transformation, which is essential for ensuring biosafety and biostability of drinking water post intrusion events.

2. Materials and methods

2.1. Location and sampling

Tap water was collected from the laboratory at the Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing, China. The water supplied to this laboratory undergoes a comprehensive treatment process, including pre-ozonation, coagulation, sedimentation, sand filtration, activated carbon filtration, and disinfection. To ensure that fresh tap water was obtained from the network, the taps were opened and allowed to run for 15 min prior to sampling, until a constant temperature was reached. Sewage samples were collected from the influent of a sewage treatment plant in Tongzhou District, Beijing, China, which primarily treats domestic sewage from urban areas. During transportation to the laboratory, samples were kept in the dark and on ice. The water quality indices can be found in Table S1.

2.2. Experimental set-up

The experiment consisted of two groups, one group is drinking water from the tap (referred as TW), the other group added 1 % volume of sewage to simulate intrusion (S-TW). In the reports of sewage intrusion incidents, there is no record of the volume/proportion of intrusion. The difference caused by the volume or concentration of intrusion is also worth investigating, but this is not the research content of this experiment. We chose a volume of 1 % because of the following reasons: (1) it is commonly used for water inoculum study (Logue et al., 2016); (2) to achieve sufficient initial cell concentrations around 10^4 cells/mL (Sousi et al., 2018); (3) it falls in the ranges that used by previous sewage intrusion simulation studies (0.01 % - 2 %)(Heibati et al., 2017; Prest et al., 2013; Vang et al., 2014). In areas with over 500,000 people, water ages are usually 3–7 days, while in areas with fewer than 50,000 people, water ages can be longer (Riley et al., 2011). A seven-day experiment was therefore constructed to simulate the dynamics of drinking water. In order to prevent contamination from dissolved organic matter leaching from the polymer pipes, glass bottles were selected as the experimental container. To simulate the worst scenario, experiment was started when residual chlorine was below 0.02 mg/L (HACH DR300), representing extreme conditions in chlorinated systems (e.g., China, U.S., Australia) and general conditions in unchlorinated systems (e.g., the Netherlands). The bottles were assembled and fixed on a shaker (80 r/min), kept in the dark with a constant temperature of 25 °C. Sampling was conducted at the same fixed time each day for chemical and microbiological analysis. On the first day, samples were collected immediately after adding 1 % sewage. In total, 42 samples were collected, which were taken from 2 groups, sampling on 7 days, each in triplicates ($n = 2 \times 7 \times 3 = 42$). Analysis was performed on dissolved organic carbon (DOC), Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR MS), total cell counts (TCC), adenosine triphosphate (ATP), 16S rRNA high throughput sequencing. The DOC, TCC, ATP measurements, DNA extraction for sequencing, and water sample solid phase extraction for FT-ICR MS were performed within 12 h after sampling. Triplicate

samples were composited for FT-ICR MS analysis ($n = 14$).

2.3. Experimental preparation

Tap water with and without sewage intrusion were prepared in 5-liter glass bottles with a customized sealing constructed, as described by Logue et al. (Logue et al., 2016). The bottles were pretreated to avoid the introduction of dissolved organic matter (DOM) contamination (Hammes and Egli, 2005). Briefly, the bottles were soaked in HCl (0.2 N) overnight, then rinsed with ultrapure water, air-dried, and heated in a muffle (450 °C) for 6 h.

The sewage as inoculum arriving at the laboratory is immediately prefiltered (150 μm; nylon mesh filter) and then filtered through 10 μm glass fiber filters (Whatman) to remove large particles and algae (Raza et al., 2022). The inoculum and collected tap water were stored in the dark at 4 °C together in properly cleaned glass bottles overnight. To avoid contamination, the metal needles, customized caps, and all glassware used in the experiments were all acid-washed overnight, extensively rinsed with Milli-Q water, and heated-sterilized at 120 °C or combusted at 450 °C for 6 h. In particular, all materials associated with microbiological indicators are sterile.

2.4. Bacterial and dissolved organic matter quantification

2.4.1. Total cell count (TCC)

Samples stained for bacterial DNA were rapidly enumerated for the total cells in the water using a flow cytometer as previously (Hammes et al., 2008; Hammes and Egli, 2005; Nescerecka et al., 2016). In short, samples (1 mL) were stained with 10 μL mL^{-1} SYBR® Green I (1:100 dilution in DMSO; Molecular Probes) and incubated in the dark for 10 min at 35 °C before measurement. Flow cytometric measurements were performed, using an Agilent NovoCyte 1040® (NovoCyte, USA).

2.4.2. Adenosine triphosphate (ATP)

Samples in 96-well plates were analyzed for ATP using GloMax® Navigator Microplate Luminometer (Promega, USA) and Water-Glo™ Detection Reagent (Promega, USA) as described previously (Yao et al., 2023). Briefly, 200 μl of the sample with 800 μl of Water-Glo™ Lysis Reagent was added to a sterile Eppendorf tube and incubated for 2 h at room temperature. Then 125 μl of each sample was taken in a 96-well plate and three times were taken as measurement replicates. The standard curve was re-measured for daily measurements to ensure accuracy. Finally, the ATP concentration of each sample was calculated from the standard curve, and the R^2 of the standard curves were all greater than 0.99.

2.4.3. Dissolved organic carbon (DOC)

15 mL water samples for DOM analysis were filtered through a 0.22 μm Whatman GF/F glass fibre membrane that had been precombusted (450 °C for 6 h). The concentration of DOM, as indicated by dissolved organic carbon (DOC), was measured using a total organic carbon (TOC) analyzer (TOC-L, Shimadzu, Japan). The detection limit and analytical precision were 0.01 mg L^{-1} and ± 2 %, respectively.

2.5. DNA extraction and sequencing

A volume of 500 mL water was filtrated through 0.22 μm filter (Whatman, UK) for DNA extraction and stored at −20 °C before extraction. DNA was extracted using the FastDNA SPIN Kit for Soil (MP Biomedicals, Santa Ana, California, USA) according to the manufacturer's protocol. For sequencing, the V3-V4 regions of 16S rRNA genes were amplified by the primers (341F: 5′-CCTACGGGNGGCWGCAG-3′ and 805R: 5′-GACTACHVGGGTATCTAATCC-3′)(Yao et al., 2023). PCR products were purified and prepared for sequencing on an Illumina Nova 6000 platform (Guangdong Magigene Biotechnology Co., Ltd) to obtain 2 × 250 bp paired end reads. The primers were removed from the generated raw sequences by employing cutadapt 3.1 (Martin, 2011). For quality filtering, error rate learning, sample inference, redundancy removal (of duplicate sequences), paired read merging, amplicon sequence variant (ASV) table construction, and chimera removal, we utilized DADA2 v1.21.0 (Callahan et al., 2016). Subsequently, we obtained an amplicon sequence variant table. Taxonomic classification and phylogenetic reconstruction were conducted through the utilization of QIIME2 2020.11 (Bolyen et al., 2019). The sequencing data have been deposited in the NCBI database, with reference code PRJNA1077455.

2.6. Extraction and FT-ICR MS analysis of DOM

The 500 mL water samples were filtered using 0.22 μm glass fiber filter membrane and acidified with HCl to pH 2 for solid phase extraction (Dittmar et al., 2008). Specifically, (1) the Agilent Bond Elut-PPL (500 mg per 6 mL) cartridges were pre-activated with 12 mL of methanol followed by acidified ultrapure water (pH 2). (2) Samples were passed through the PPL cartridges at a flow rate of 5 mL/min. (3) The cartridges were rinsed with 12 mL of acidified ultrapure water (pH 2) to remove salts and then dried with nitrogen gas. (4) The cartridges were eluted with 12 mL of methanol. (5) The eluted samples were blow-dried with nitrogen, redissolved in 1 mL of methanol, and stored in the dark at −20 °C. Highly accurate measurement of the molecular composition of DOM in water samples using 15.0 Tesla Bruker Solarix FT-ICR MS system (Bruker Daltonics, Billerica, MA) equipped with negative electrospray ionization (ESI) ion source. Details on analytical conditions and molecular formula assignments are provided in the supplementary information. The double bond equivalence minus oxygen (DBE-O), the modified aromaticity index (AI_{mod}) and the nominal oxidation state of carbon (NOSC) of each compound were calculated as follows:

$$DBE - O = 1 + 0.5(2C - H + N) - O \quad (1)$$

$$AI_{mod} = \frac{1 + C - 0.5O - S - 0.5(N + H)}{C - 0.5O - S - N} \quad (2)$$

$$NOSC = 4 - \frac{4C + H - 3N - 2O - 2S}{C} \quad (3)$$

Where C, H, O, N, and S refer to the number of atoms per formula of carbon, hydrogen, nitrogen, oxygen, and sulfur, respectively. The elemental combinations were CHO, CHON, CHOS and CHONS. Meanwhile, five classes of compounds were identified based on the criteria of AI_{mod} and H/C (Hou et al., 2022; Kellerman et al., 2014): polycyclic aromatics (PA, $AI_{mod} > 0.66$), polyphenols (Poly, $0.50 < AI_{mod} \leq 0.66$), highly unsaturated and phenolic compounds (HuPh, $AI_{mod} \leq 0.50$, $H/C < 1.5$), aliphatic compounds (Ali, $1.5 \leq H/C \leq 2.0$) and carbohydrate compounds (Carbo, $2.0 < H/C$).

2.7. Statistical analysis

2.7.1. Alpha and beta diversity

The alpha diversity of bacteria was calculated using ASV number, Chao1, Shannon-Wiener (Shannon), and Gini-Simpson (Simpson) indices. The relative peak intensities of molecules were calculated by normalizing signal intensities of assigned peaks to the sum of all intensities within each sample. The alpha diversity of DOM based on relative peak intensities was calculated using molecule number, Shannon-Wiener (Shannon) and Gini-Simpson (Simpson) indices. Non-metric multidimensional scaling (NMDS) was performed based on the Bray-Curtis dissimilarity metric of bacteria or DOM, and then the effect of inoculation on bacterial or DOM composition of distance metrics was statistically tested via ADONIS. These analyses of bacterial and chemical diversity were conducted using vegan package (Dixon, 2003).

2.7.2. Cumulative biomass and DOM

Only a small proportion of DOM is available to bacteria (BAC) as an

energy source for growth (Prest et al., 2016), so it is difficult to directly observe changes in the daily relative intensities of all DOM molecules (Figure S1). Given that the daily changes in abundance of microbially affected DOM should be correlated with the biomass for that day, we proposed the metric of cumulative biomass to identify the changes of DOM molecules affected by microbes.

$$\Delta DOM_{day_n} = k \times BAC_{day_n} \quad (4)$$

$$DOM_{day_n} = DOM_{day_0} - \sum_{i=0}^{n} \Delta DOM_{day_i} = DOM_{day_0} - \sum_{i=0}^{n} k \times BAC_{day_i} \quad (5)$$

According to this concept, cumulative TCC and cumulative ATP were calculated. From the linear regression of cumulative TCC and cumulative ATP (Figure S2), it can be deduced that their Spearman correlations with the DOM molecules are congruent. Therefore, cumulative TCC was used as the outcome of the Spearman correlation with DOM molecular relative peak intensities.

2.7.3. Bacterial ASVs vs. DOM molecules

To link the bacterial ASVs and DOM molecules, we followed the procedure described by Osterholz et al. (Osterholz et al., 2016). Briefly, the subsets of principal coordinates including at least 80 % of the total variation of each data set were selected (DOM: PCoA1–3 of TW, PCoA1–2 of S-TW; Bacteria: PCoA1–2 of TW or S-TW). Then, a symmetric canonical correlation analysis (CCorA) was performed to identify the first canonical axis pair representing the highest correlational dimensions in bacterial and DOM space, to account for their bidirectional interactions. The asymmetric redundancy analysis (RDA) was utilized to establish the explanatory capacity of bacteria on DOM, and vice versa. Furthermore, we calculated the Spearman rank correlations between the relative peak intensities with the first canonical axis in DOM space. Similarly, we calculated the Spearman rank correlations between relative abundance with the canonical axis in bacterial space. To investigate the primary impacts of bacteria and DOM covariation, linear regression analyses of the first canonical axis pair on bacterial and DOM richness and evenness diversity (Shannon and Simpson) were performed in the two groups.

2.7.4. Co-occurrence network

Co-occurrence network analyses were performed to investigate the relationship between microbial ASVs and DOM molecules. The analyses used the relative abundance of ASVs and the relative peak intensity of molecules Spearman correlation results. Rare ASVs or molecules (relative abundance < 0.01 %) were excluded before correlation analysis. For statistical significance, only microbial ASVs or DOM molecules present in more than half of the samples were included in the Spearman correlation analysis. To reduce statistical errors, all p-values were adjusted using the Benjamini-Hochberg method (Benjamini and Hochberg, 1995). Bilateral networks were constructed using robust correlations ($|\rho| \geq 0.7$ and $P < 0.05$). Network visualization and module partitioning were performed using Gephi (https://gephi.org). Each network node was assigned roles based on the computation of intra-module connectivity (Zi) and inter-module connectivity (Pi) using the R package of igraph. Nodes have been divided into four categories (Guimera and Nunes Amaral, 2005): (1) peripheral nodes (Zi ≤ 2.5, Pi ≤0.62), (2) connectors (Zi ≤ 2.5, Pi >0.62), (3) module hubs (Zi > 2.5, Pi ≤0.62), and (4) network hubs (Zi > 2.5, Pi >0.62). All the statistical analyses were performed in R (version 4.3.1, R Core Team, 2023).

3. Result

3.1. Bacterial growth and communities' succession

Despite sewage intrusion accounting for only 1 %, it led to an immediate 8-fold increase in total cell count (TCC) (Fig. 1a, 7.8×10^4 cells/mL) compared to uncontaminated tap water. In the tap water, cell concentration remained stable until an observed increase from Day-3 to Day-6, with a growth rate of 1.9×10^4 cells/mL•day. In contrast, after sewage intrusion, the bacterial growth rate surged to as high as 8.6×10^4 cells/mL•day, reaching its peak value of 3.4×10^5 cells/mL on Day-3, then gradually decreasing and stabilizing at 2.3×10^5 cells/mL on Day-5 and Day-6. The active biomass, measured by ATP, exhibited a similar trend to TCC (Figure S3a).

For the bacterial community, the intrusion of 1 % sewage elevated the number of observed Amplicon Sequence Variants (ASVs) by 459, which increased from 939 in tap water to 1398 ASVs post-intrusion

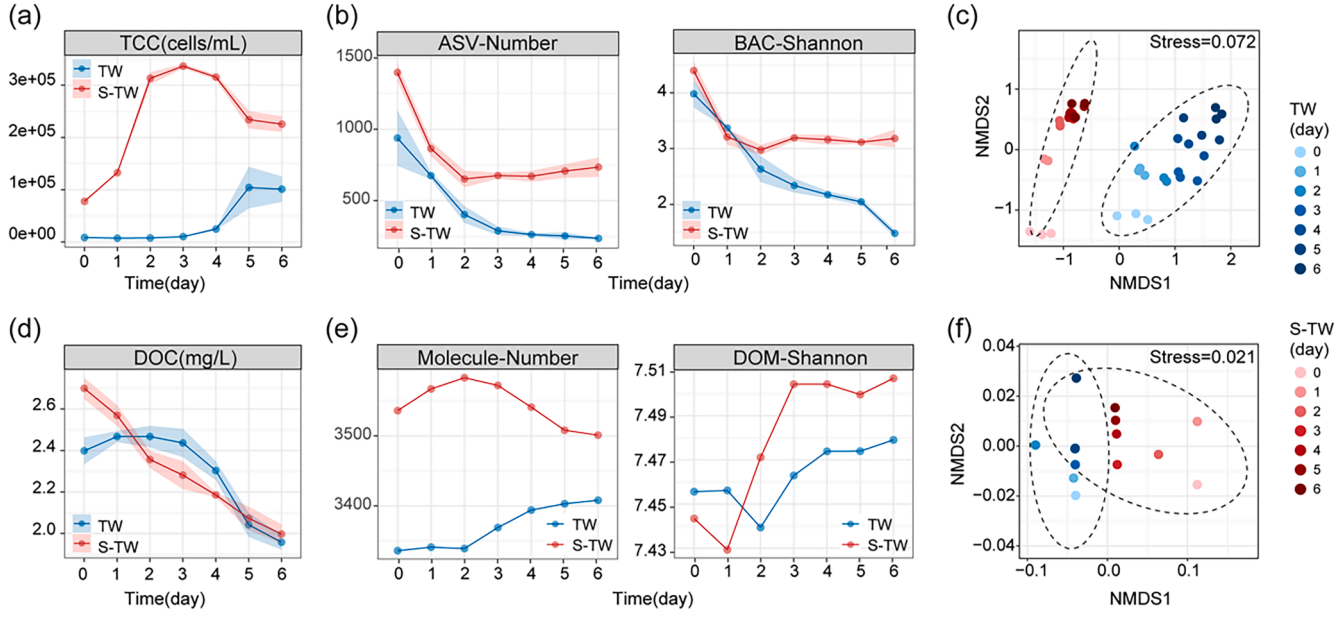


Fig. 1. Seven days' dynamic of bacteria and dissolved organic matter. (a) changes of biomass quantified by total cell count (TCC); (b) alpha diversity of bacterial community; (c) NMDS plots based on Bray-Curtis dissimilarity of bacterial community; (d) changes of dissolved organic carbon concentrations (DOC); (e) alpha diversity of DOM; (f) NMDS plots based on Bray-Curtis dissimilarity of DOM. Lines with ribbons in subfigures (a, b, d) indicate means ± standard deviation (SD) ($n = 3$).

(Fig. 1b). On average, sewage intrusion introduced 876 ASVs. The number of sewage intrusion introduced ASVs was higher than the ASV number differences between TW and S-TW (876 vs. 459). This is because a considerable amount of ASVs in TW were not detected after sewage intrusion (Venn diagram, Figure S3d), which is likely due to their too low abundance to be detected by sequencing after 8-fold increase in biomass. In contrast to biomass development, the observed ASV numbers in tap water, with or without sewage intrusion, both sharply decreased in the first three days and remained stable afterward. The sewage intrusion group consistently exhibited higher alpha diversity than tap water (ASV number, Fig. 1b). Similar trends were observed for other alpha diversity indexes, e.g., Chao1, Shannon and Simpson (Figure S3). Combining the increase in biomass and the decrease in observed ASV number, it is reasonable to hypothesize that certain species became dominant during the growth. As revealed by the NMDS plot, there were significant differences in the bacterial communities of tap water with and without sewage intrusion ($P < 0.001$, ADONIS, Fig. 1c), indicating that the intrusion of sewage might have changed the dominant members in the bacterial community.

At the phylum level, the dominance of Proteobacteria (52 %) and Actinobacteriota (20 %) in tap water was replaced by Campylobacteriota (50 %) and Bacteroidota (15 %) after the intrusion of 1 % sewage. Over time, the relative abundances of Proteobacteria increased in both cases, rising to 80 % at Day-1 and over 90 % after Day-2 in tap water, while increasing from 10 % to 70 % at Day-1 and further to 90 % after Day-3 in sewage-introduced tap water (Figure S4a). However, the results at the class level indicated that Alphaproteobacteria governed tap water (64 %), whereas sewage-introduced samples were governed by Gammaproteobacteria (84 %) on Day-6 (Figure S4b). At the genus level, *Porphyrobacter* spp. and *Nevskia* spp. became dominant in tap water on Day-4 and Day-5, corresponding to the growth curve of TCC shown in Fig. 1a. In sewage-contaminated tap water, the dominance of *Arcobacter* spp. and *Pseudarcobacter* spp. was replaced by *Acinetobacteria* spp. (69 %) from Day-1 onwards (Figure S4e). However, it is noteworthy that sewage introduced 300 ASVs belonging to pathogenic genera (Figure S5), such as *Acinetobacter* spp., *Arcobacter* spp., *Aeromonas* spp., *Pseudomonas* spp., and *Legionella* spp., confirming the potential microbial risks. Moreover, the growth dynamic results showed that most of the introduced ASVs (89 %) disappeared shortly after the intrusion. Some ASVs (11 %), however, were able to survive and grow in the new environment, posing higher and longer risks in drinking water systems. Specifically, five ASVs in *Pseudomonas* spp. showed a significant increase in relative abundance.

3.2. Concentration and composition of dissolved organic matter

Quantitatively, the intrusion of 1 % sewage resulted in a 12.5 % increase in the concentration of dissolved organic carbon (DOC) from 2.4 mg/L to 2.7 mg/L (Fig. 1d). Over time, the DOC in tap water did not decrease until Day-4 when cell numbers began to increase. In contrast, after sewage intrusion, the trend shifted to a continuous decrease, even when cell numbers started to decline after Day-4. Qualitatively, the intrusion of 1 % sewage introduced 452 unique molecular formulae, accounting for 13.5 % of the originally present formulae in tap water (3336, Figure 1e; Venn diagram, Figure S3f). The newly introduced DOM molecules were predominantly CHOS (63.7 %), followed by CHON (22.1 %), CHO (10.8 %), and CHONS (3.3 %). These differences contributed to significant variations in beta diversity but not Shannon diversity between tap water with and without sewage intrusion (Figure 1e; $P < 0.001$, ADONIS, Fig. 1f).

Interestingly, although DOC consistently decreased in both cases (Fig. 1d), the number of molecular formulae for DOM changed with a different trend, aligning well with the dynamics of Total Cell Count (TCC). In brief, the number of detected molecular formulae remained stable until Day-3 and started to increase until Day-6 in tap water, while it increased immediately from Day-0 to Day-3 and subsequently decreased until Day-6 in sewage-contaminated tap water. The well-fitted trends between biomass and DOM diversity suggest possible correlations between microbial activity and community with the concentration and composition of DOM. However, bacterial metabolism did not significantly contribute to the alpha diversity of DOM (variation within the range of 0.1), possibly because bacterial growth only involved a small portion of DOM molecules (15–25 % DOC).

3.3. DOM transformation mediated by bacterial metabolism

Focusing on the dynamics of the bacterial-related portion of DOM, molecules that positively correlated (generated by bacteria) and negatively correlated (consumed by bacteria) with cumulative biomass throughout the study period were depicted in Van Krevelen (VK) plots (Spearman's correlation coefficient $|\rho| \geq 0.6$), as shown in Fig. 2.

The number of identified molecules associated with bacterial metabolism sharply increased following sewage intrusion, rising from 750 molecules (323 positive, 427 negative) to 1798 molecules (1321 positive, 477 negative) (Fig. 2). Notably, there were 998 more molecules generated (positively related) from bacterial metabolism after sewage intrusion. In both cases, the generated molecules exhibited moderate H/C and O/C ratios, concentrated in the middle of the VK plot (Fig. 2b). The number of bacteria-consumed molecules (negatively related) were similar. In tap water, bacteria consumed molecules with low H/C and high O/C ratios, high H/C and low O/C ratios, and moderate H/C and O/C ratios. After sewage intrusion, bacteria also consumed molecules with either low H/C and high O/C or high H/C and low O/C ratios but left the moderate H/C and O/C molecules unused. Moreover, there is little overlap of molecules between tap water with and without sewage intrusion, considering either bacteria-consumed (54 molecules, 6.4 %) or generated molecules (141 molecules, 9.4 %) (Figure S6). These differences in molecules may stem from different dominant bacteria consuming and generating distinct molecules, or from the contribution of newly introduced molecules and their transformation products. Specifically, 242 out of 452 sewage-introduced molecules were consumed by bacteria, predominantly CHOS molecules with high H/C and low O/C ratios (Figs. 2, 3a, Figure S7).

Furthermore, considering DBE-O, AI_{mod}, and NOSC, the bacterial generated molecules were generally lower than inherent molecules consumed by bacteria in both cases, the differences of which were all significant except for NOSC in tap water (Fig. 3a). The portion of sewage-introduced molecules consumed by bacteria exhibited even lower AI_{mod} and NOSC than the bacteria-generated molecules. Regarding elemental composition, the bacterial-consumed molecules were predominantly CHO, whereas the bacterial-generated molecules contained less CHO but more CHON than the bacterial-consumed ones (Fig. 3b). In terms of compound composition, it is noteworthy that all bacterial metabolism-related molecules were dominated by HuPh (Fig. 3c). The bacterial-consumed compounds were more diverse than the bacterial-generated ones in both cases. Specifically, Carbo and Poly were converted into HuPh or Ali. Importantly, the sewage-introduced DOM that was consumed by bacteria had a different compound composition than inherent compounds, being dominated by Ali (53.3 %) and HuPh (41.7 %).

3.4. Covariation of bacterial community and DOM composition

Statistical analysis between bacterial and DOM compositions was conducted at six phylogenetic levels (Table 1, from phylum to ASV). According to canonical correlation analysis (CCorA), a robust correlation was observed between changes in the bacterial community and the DOM molecular composition in tap water, both with and without sewage intrusion. This correlation was most significant at the ASV and genus levels, diminishing with lower phylogenetic levels and becoming statistically insignificant at the phylum level. Consequently, the ASV level was selected for further analysis of the relationships between

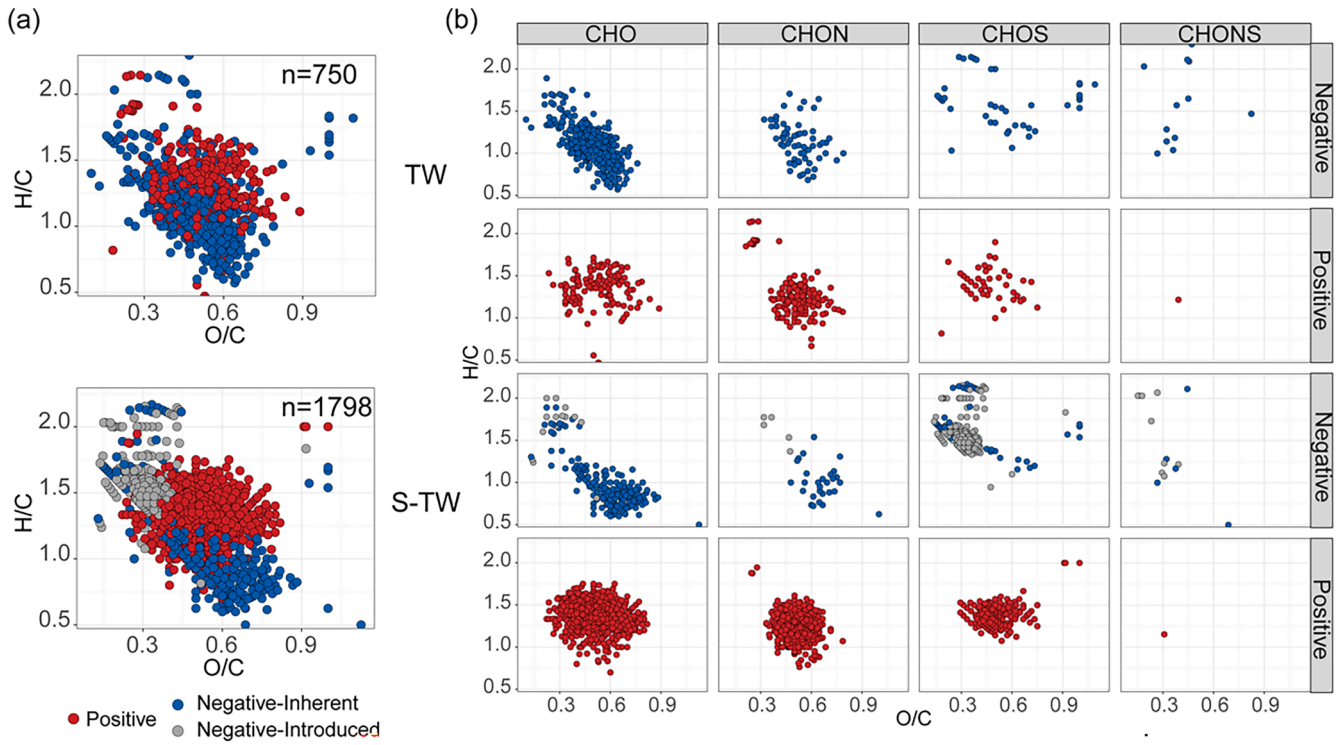


Fig. 2. Van Krevelen diagrams of DOM molecules classified as positively and negatively correlated with cumulative biomass. (a) overall molecules marked as positively correlated (bacterial generated, in red) and negatively correlated (bacterial consumed, in grey and blue) in tap water (TW) and tap water with 1 % sewage intrusion (S-TW); (b) DOM molecules grouped into different elemental compositions. Positively correlated molecules (bacterial generated) marked in red, negatively correlated molecules (bacterial consumed) marked in grey (sewage introduced molecules) and blue (tap water inherent molecules).

bacterial community members and DOM molecules.

Furthermore, redundancy analysis (RDA) results indicated that the explanatory power of DOM composition for variations in the bacterial community was greater than that of the bacterial community for variations in DOM composition in tap water. Conversely, in tap water with sewage intrusion, the bacterial community exhibited a greater potential for explaining changes in DOM composition than the explanatory power of DOM composition on the bacterial community. This suggests that the intrusion of sewage altered the driving force in the covariance of bacterial community and DOM composition. As confirmed by regression analyses of the covariance dimension, the dominance of DOM selective effects on shaping the bacterial community in tap water shifted to the governance of bacterial metabolism on the molecular composition of DOM (Figure S9).

The DOM molecules and bacterial ASVs associated with the first canonical axis were identified through Spearman correlation and are presented in Fig. 4 (Spearman's correlation coefficient $|\rho| \geq 0.6$). Interestingly, the selected DOM molecules were nearly identical to those shown in Fig. 2, chosen due to their strong correlations with cumulative biomass. Specifically, the molecules identified by both approaches were entirely the same for tap water, and for tap water with sewage intrusion, >93 % overlapped. This provides direct evidence that, during its distribution, not only biomass but also the bacterial community plays a crucial role in the transformation of DOM in drinking water. Concerning the bacterial community, the majority of ASVs strongly correlated with the first canonical axis belonging to Proteobacteria. At the class level, the most positively correlated ASVs in tap water were predominantly Alphaproteobacteria, while in tap water with sewage intrusion, they were a combination of Alphaproteobacteria and Gammaproteobacteria (Fig. 4b).

3.5. Interactions between bacterial ASVs and DOM molecules

Identified through co-occurrence network analysis, there were 117 ASVs and 108 molecules in tap water, which increased to 293 ASVs and 945 molecules after sewage intrusion (Fig. 5a). The intrusion of sewage significantly amplified the complexity of co-occurrence networks, expanding from 225 nodes and 310 edges in tap water to 1238 nodes and 9094 edges in tap water with sewage intrusion, both exhibiting a well-defined modular structure (modularity > 0.4). Based on connectivity within and among modules, 0 network hubs, 7 module hubs, 26 connectors, and 192 peripherals were identified in the tap water network, which increased to 6 network hubs, 50 module hubs, 179 connectors, and 1003 peripherals in tap water with sewage intrusion (Fig. 5b). Notably, module hubs were primarily DOM molecules (71 %) in tap water, whereas after sewage intrusion, the major network and module hubs shifted to bacterial ASVs (89 %) (Table S4). For the top 7 major modules of each network (Fig. 5c), Proteobacteria (54 %) accounted for the largest proportion in tap water, with three modules exclusively consisting of this phylum. However, the intrusion of sewage introduced more diverse phyla, predominantly dominated by Proteobacteria (42 %) and Bacteroidota (29 %). Regarding DOM, HuPh (93 %) accounted for the largest proportion in the tap water network, with most modules (5/7, modules 2–6) consisting exclusively of HuPh. Following sewage intrusion, all seven modules displayed a more diverse and similar DOM composition, dominated by highly HuPh (82 %) and Ali (16 %).

4. Discussion

Combining the advancements in analytical chemistry and molecular biology, this study investigated the immediate contamination and long-term disruptive effects of sewage intrusion into drinking water. Going beyond prior research, this study sheds light on the intricate dynamics of bacterial communities and DOM molecules following sewage intrusion.

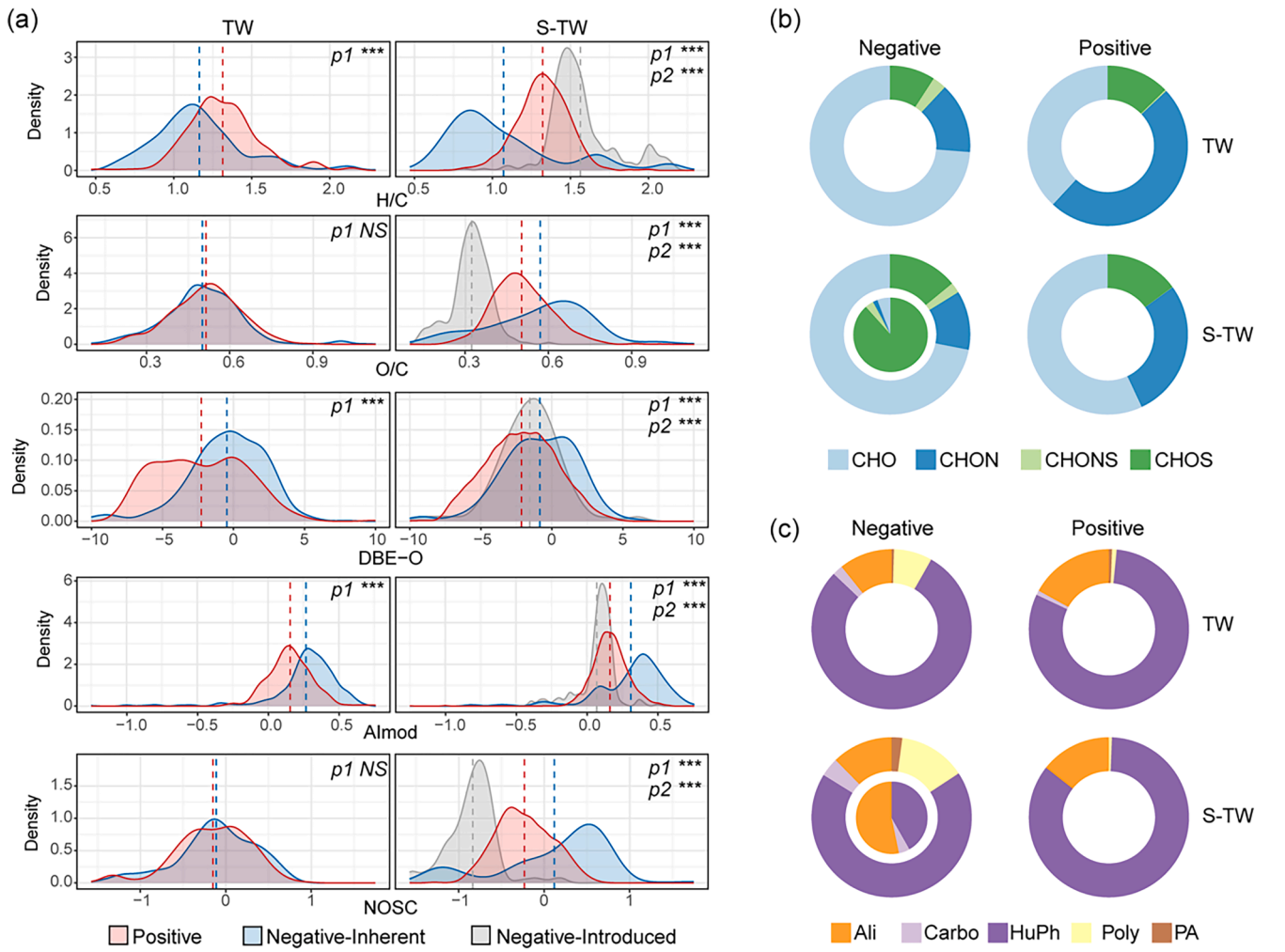


Fig. 3. Chemical properties and compositions of DOM molecules. (a) Density plots of the distribution and mean of DOM molecules on H/C ratios, O/C ratios, DBE-O, AI_{mod}, and NOSC. Two-tailed Wilcoxon rank-sum test for comparing significant differences in molecular traits. Significance: ***, $P < 0.001$; Generated molecules (positively correlated) in Red, consumed molecules (negatively correlated) in Gray (inherent molecules) and Blue (sewage introduced molecules). (b) Elemental composition of bacterial generated and consumed DOM molecules. (c) Compounds composition of bacterial generated and consumed DOM molecules. In (b) and (c), sewage introduced molecules that consumed by bacteria are presented in the inner circle.

Table 1
The correlations between DOM and bacteria at different phylogenetic levels.

	Phylum	Class	Order	Family	Genus	ASV
(a) CCorA coefficients: the CanAxis1						
TW	0.98	0.99*	0.98*	0.98*	0.99**	0.99**
S-TW	0.90	0.92**	0.93**	0.94**	0.94**	0.94**
(b) Redundancy coefficients: BAC on DOM / DOM on BAC						
TW	0.31/0.94	0.27/0.68	0.29/0.84*	0.37*/0.79*	0.46**/0.78**	0.46**/0.78**
S-TW	0.31/0.67*	0.74**/0.66*	0.76***/0.68***	0.80**/0.68**	0.80**/0.69**	0.80**/0.69***

Provided are (a) the *first canonical axis* correlation coefficients resulting from CCorA and global significances (permutational probability), and (b) adjusted redundancy coefficients and significance for DOM composition as explained by bacterial composition (and vice versa). DOM: dissolved organic matter; BAC: bacteria; significance:
* $P < 0.05$.
** $P < 0.01$.
*** $P < 0.001$.

The finding of this work adds new dimension on understanding and assessing the risks and impacts of sewage intrusion into drinking water systems.

4.1. *Immediate contamination following sewage intrusion*

Considerable quantitative and qualitative contaminations regarding bacteria and DOM were immediately detected after intrusion of 1 % sewage. This contamination was observed as 8-fold increase in biomass (TCC), 48.9 % increase in bacterial species (ASVs), 12.5 % increase in organic carbon content (DOC), and 13.5 % increase in unique DOM molecular formulae. These findings align with previous studies that have detected sewage introduced microbial contamination either by ATP/TCC based bacterial quantification (Vang et al., 2014) or

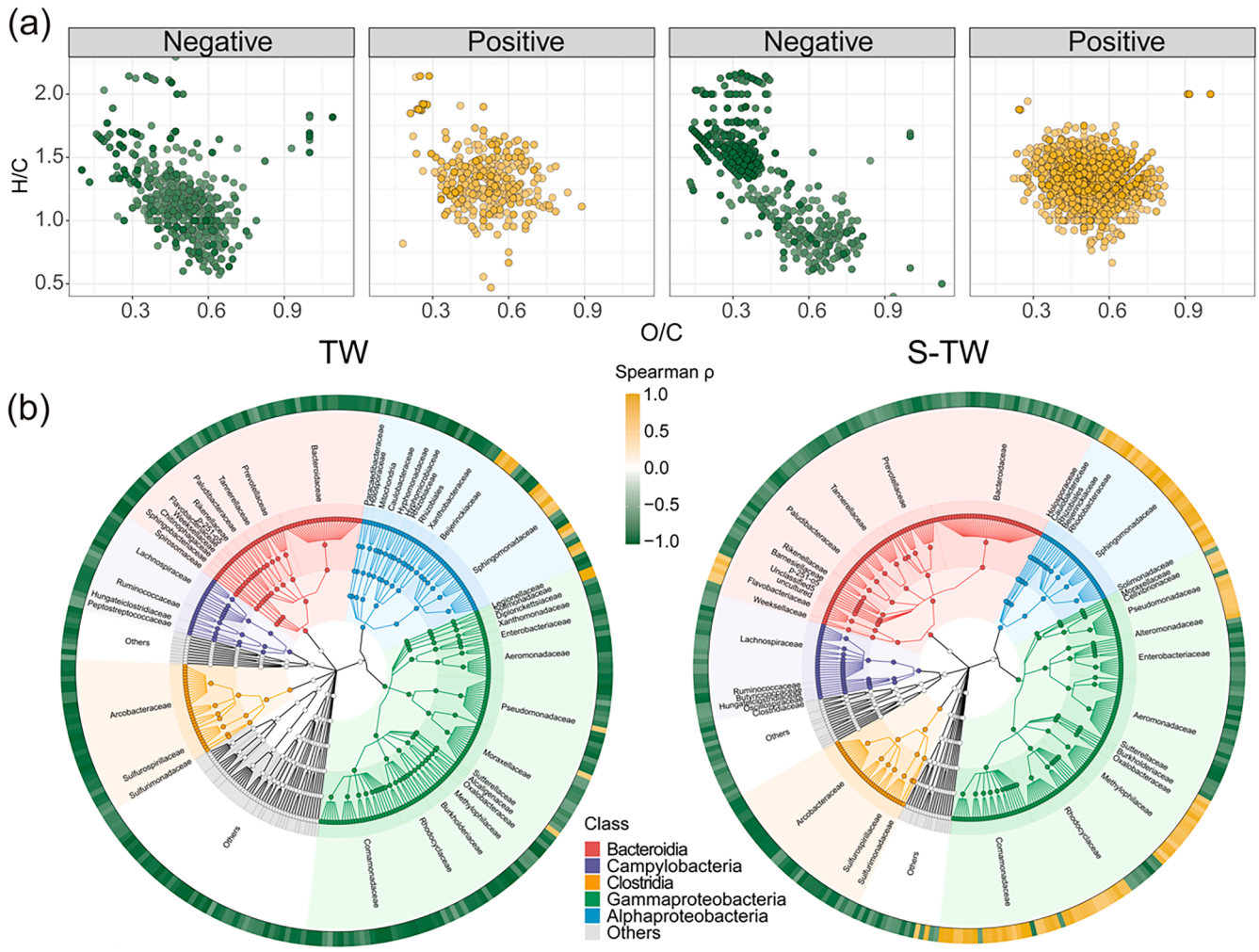


Fig. 4. Covaried DOM molecules and bacterial ASVs. (a) Van Krevelen plots show the correlations of molecular formulae with the first canonical axis associating DOM composition with bacterial community composition. Plotting was ordered by the absolute correlation value. (b) Taxonomic cladograms show the ASVs (relative abundance > 0.01 %) correlated with the first canonical axis associating ASVs with DOM composition. Family-level names are labeled on the taxonomic cladograms. Use identical colors (Gold and green) to show the intensity of Spearman's correlation coefficient ρ of molecules or ASVs.

fluorescence-based DOM characterization (Sorensen et al., 2021), which could detect 10^2 - 10^4 times diluted sewage contamination. With the primary focus of the potential health risks associated with human pathogens, online monitoring methods and devices have been developed and validated for the fast detection and early warning of possible sewage intrusion into drinking water system (Favere et al., 2021; Sorensen et al., 2018; Stedmon et al., 2011). Beyond previous reports, this study characterized the changes in bacterial community and DOM composition following sewage intrusion. It was found that the intrusion of 1 % sewage introduced 876 bacterial ASVs and 452 DOM molecules into drinking water. Among the introduced bacterial species, 300 ASVs assigned to pathogen containing genera, confirming the widely reported potential risk to consumers' health (Craun et al., 2010; Ercumen et al., 2014; Säve-Söderbergh et al., 2017). Besides the immediate infection risks, the introduced species that survived and multiplied (43/300 ASVs, including 5 ASVs assigned to *Pseudomonas* spp.) in drinking water environment may pose even higher risks to customers. Moreover, the high number and concentration of introduced DOM molecules suggest that in-situ fluorescence-based spectroscopy may be a rapid and resilient approach for monitoring and detecting sewage contamination in drinking water systems (Gunter et al., 2023).

4.2. *Profound disruptive effects induced by sewage intrusion*

The intruded bacterial ASVs and DOM molecules will stay in drinking water system before they are flushed out at taps. Therefore, they may have more long-lasting impacts on drinking water microbiology than immediate contamination following intrusion. By this seven-days dynamic study, we have observed profound effects induced by sewage intrusion regarding the growth dynamic, community succession, and metabolic products. Specifically, sewage intrusion greatly elevated bacterial growth (5 folds), selectively promoted dominant ASVs, and sharply increased metabolite diversity (998 more unique molecules). More profoundly, according to statistical analysis, the intrusion of 1 % sewage shifted the driving force of bacterial ASV and DOM molecules covariance from DOM dependent process in tap water to bacterial governed process. In other words, the sewage intrusion played a key role in developing new microbial ecology, which shifted bacterial community and metabolism. This may be explained by the relationship between bacterial community and DOM composition, such as specific carbon compounds triggered growth of specific bacterial species, and in return, metabolism of different bacterial species produced diverse metabolites (Cottrell and Kirchman, 2000; Osterholz et al., 2016; Yang et al., 2020a). The residential environment of oligotrophic drinking water with low biomass, and the high quantity and diversity of bacterial ASVs and DOM molecules in sewage may favor the success of invaders after sewage

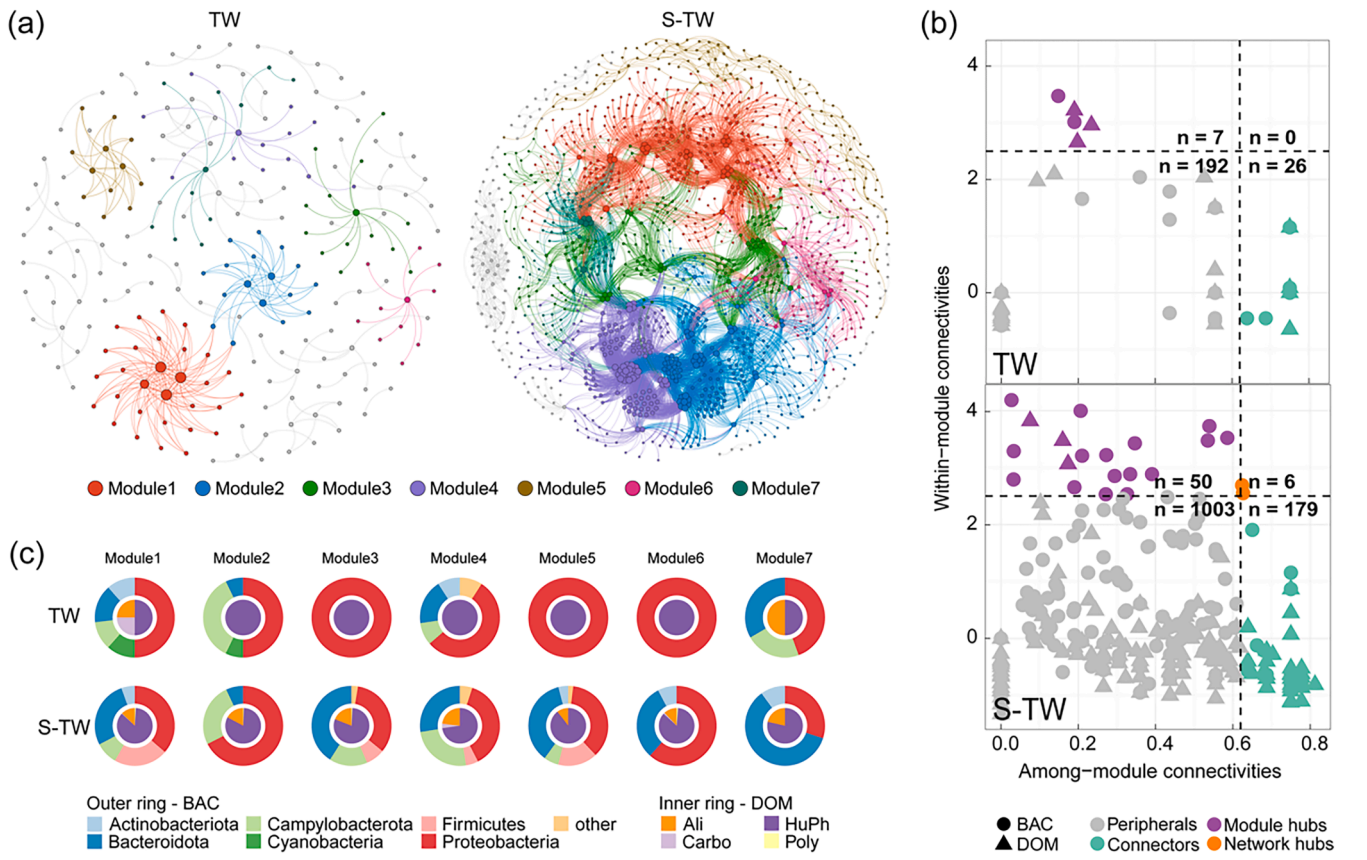


Fig. 5. Co-occurrence network between bacterial ASVs and DOM molecules. (a) The topology of the networks. Network nodes represent DOM molecules or bacterial ASVs; the top seven modules are shown in different colors. (b) Node importance identified by within-module and among-module connectivity. The detail information of bacterial ASVs and DOM molecules is given in Table S4. (c) Associations of bacterial ASVs and DOM molecules in different modules.

intrusion (De Roy et al., 2013; Kinnunen et al., 2016; van Elsas et al., 2012; Vila et al., 2019). However, the experimental design of our study does not allow us resolving molecular mechanism on the microbial ecology development; rather, we provide new dimension of insights on the covariance of bacterial community and DOM composition, especially the disruptive changes following sewage intrusion.

4.3. *Implications and outlook*

This study is among the first to uncover the dynamic changes regarding bacterial community and DOM molecules in drinking water after sewage intrusion. It is clear that sewage intrusion induced more disruptive effects on microbial ecology rather than only on the widely reported immediate health risk. This highlighted the importance of long-term thinking in monitoring and managing risks following sewage intrusion into drinking water. Efforts have been made to develop methods and devices for the rapid detection of sewage intrusion and other microbial contamination in drinking water distribution system, which is undoubtably important, especially considering the potential threats to human health (LeChevallier et al., 2003; Xu et al., 2019). However, until now, there have been no guidelines on how to monitor and manage the long-term risks following sewage intrusion. Based on our results, it is highly recommended that environmental agencies and water utilities pay attention to the post intrusion risks management, even without immediate warning of outbreak potentials.

It should be mentioned that there are limitations in the current simulation study and the observations in field drinking water distribution system subjected to sewage intrusion might be different, especially considering the complexity of field distribution system with various factors. For example, residual chlorine was not considered in the current study for simulating worst scenario of chlorinated systems and general situation of unchlorinated systems. However, most of the cases, residual chlorine may present during the sewage intrusion, the presence of which would place additional selection pressure on bacterial growth, community succession, and DOM transformation (Mitch et al., 2023; Wang et al., 2014; Yang et al., 2008). Additionally, different concentrations of residual chlorine may also selectively affect the communities that survive after intrusion (Luo et al., 2021). Regarding DOM, the presence of residual chlorine may react with introduced DOM causing rapid decay of chlorine while generating additional assimilable organic carbon (AOC) causing growth of both indigenous and invasive microbes (Ramseier et al., 2011). Moreover, the more CHON compounds generated by bacterial metabolism after sewage intrusion may lead to formation of more nitrogenous disinfection byproducts with higher toxicity (Mitch et al., 2023). Besides chlorine, the complexity of field distribution networks is increased by the formation and accumulation of biofilms and sediments over time (Liu et al., 2014), as well as the use of different pipe materials. When intrusion occurs with the accompanied pipe breaks and disturbances, the organics and microbes harbored by distribution pipes may be destabilized and released into water, which will add up another dimension of complexity. Further studies are recommended to address these uncertainties about sewage intrusion into drinking water distribution systems. More seriously, there might be higher risks if the intruded (opportunistic) pathogens attached and multiplied in biofilm. If so, contaminated biofilm and distribution system will keep releasing pathogens into drinking water, because biofilm may protect invaders from disinfectant and retained in distribution system much longer than the dynamically flowing water (Liu et al., 2016). The cleaning of distribution network biofilms is not only difficult but also expensive. Therefore, further research is recommended to explore the effects of

sewage intrusion on drinking water biofilm.

CRediT authorship contribution statement

Mengqing Fan: Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Conceptualization. **Anran Ren:** Visualization, Methodology. **Mingchen Yao:** Visualization, Methodology. **Xiaoming Li:** Supervision, Methodology, Conceptualization. **Walter van der Meer:** Writing – review & editing, Supervision. **Guo Yu:** Writing – review & editing, Supervision, Methodology. **Gertjan Medema:** Writing – review & editing, Supervision. **Joan Rose:** Writing – review & editing, Supervision. **Gang Liu:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Conceptualization.

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.

Acknowledgements

The authors would like to acknowledge the support from National Natural Science Foundation of China (Basic Science Center Project, 52388101; 52370105).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.watres.2024.122281.

References

Agency, U.S.E.U.S.E.P. 2011. Aging Water Infrastructure Research: Science and Engineering for a Sustainable Future. .

Antony, R., Willoughby, A.S., Grannas, A.M., Catanzano, V., Sleighter, R.L., Thamban, M., Hatcher, P.G., Nair, S., 2017. Molecular insights on dissolved organic matter transformation by supraglacial microbial communities. Environ. Sci. Technol. 51 (8), 4328–4337.

ASCE. 2022. Water Infrastructure|American Society of Civil Engineers (ASCE)'s 2021 Infrastructure Report Card.

Barton, N.A., Farewell, T.S., Hallett, S.H., Acland, T.F., 2019. Improving pipe failure predictions: factors affecting pipe failure in drinking water networks. Water Res. 164, 114926.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B (Methodological) 57 (1), 289–300.

Besner, M.C., Broseus, R., Lavoie, J., Giovanni, G.D., Payment, P., Prevost, M., 2010. Pressure monitoring and characterization of external sources of contamination at the site of the payment drinking water epidemiological studies. Environ. Sci. Technol. 44 (1), 269–277.

Besner, M.C., Prevost, M., Regli, S., 2011. Assessing the public health risk of microbial intrusion events in distribution systems: conceptual model, available data, and challenges. Water Res. 45 (3), 961–979.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J.E., Bittinger, K., Brejnrod, A., Brislawn, C.J., Brown, C.T., Callahan, B.J., Caraballo-Rodriguez, A.M., Chase, J., Cope, E.K., Da Silva, R., Diener, C., Dorrestein, P.C., Douglas, G.M., Durall, D.M., Duvallet, C., Edwardson, C.F., Ernst, M., Estaki, M., Fouquier, J., Gauglitz, J.M., Gibbons, S.M., Gibson, D.L., Gonzalez, A., Gorlick, K., Guo, J., Hillmann, B., Holmes, S., Holste, H., Huttenhower, C., Huttley, G.A., Janssen, S., Jarmusch, A.K., Jiang, L., Kaehler, B.D., Kang, K.B., Keefe, C.R., Keim, P., Kelley, S.T., Knights, D., Koester, I., Kosciolek, T., Kreps, J., Langille, M.G.I., Lee, J., Ley, R., Liu, Y.X., Loftfield, E., Lozupone, C., Maher, M., Marotz, C., Martin, B.D., McDonald, D., McIver, L.J., Melnik, A.V., Metcalf, J.L., Morgan, S.C., Morton, J.T., Naimey, A.T., Navas-Molina, J.A., Nothias, L.F., Orchanian, S.B., Pearson, T., Peoples, S.L., Petras, D., Preuss, M.L., Pruesse, E., Rasmussen, L.B., Rivers, A., Robeson, M.S., 2nd, Rosenthal, P., Segata, N., Shaffer, M., Shiffer, A., Sinha, R., Song, S.J., Spear, J.R., Swafford, A.D., Thompson, L.R., Torres, P.J., Trinh, P., Tripathi, A., Turnbaugh, P.J., Ul-Hasan, S., van der Hooft, J.J.J., Vargas, F., Vazquez-Baeza, Y., Vogtmann, E., von Hippel, M., Walters, W., Wan, Y., Wang, M., Warren, J., Weber, K.C., Williamson, C.H.D., Willis, A.D., Xu, Z.Z., Zaneveld, J.R., Zhang, Y., Zhu, Q., Knight, R., Caporaso, J.G., 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat. Biotechnol. 37 (8), 852–857.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J., Holmes, S.P., 2016. DADA2: high-resolution sample inference from Illumina amplicon data. Nat. Methods 13 (7), 581–583.

Cho, E.J., Yang, J.Y., Lee, E.S., Kim, S.C., Cha, S.Y., Kim, S.T., Lee, M.H., Han, S.H., Park, Y.S., 2013. A waterborne outbreak and detection of cryptosporidium oocysts in drinking water of an older high-rise apartment complex in seoul. Korean J. Parasitol. 51 (4), 461–466.

Cottrell, M.T., Kirchman, D.L., 2000. Natural assemblages of marine proteobacteria and members of the Cytophaga-Flavobacter cluster consuming low- and high-molecular-weight dissolved organic matter. Appl. Environ. Microbiol. 66 (4), 1692–1697.

Craun, G.F., Brunkard, J.M., Yoder, J.S., Roberts, V.A., Carpenter, J., Wade, T., Calderon, R.L., Roberts, J.M., Beach, M.J., Roy, S.L., 2010. Causes of outbreaks associated with drinking water in the United States from 1971 to 2006. Clin. Microbiol. Rev. 23 (3), 507–528.

De Roy, K., Marzorati, M., Negroni, A., Thas, O., Balloi, A., Fava, F., Verstraete, W., Daffonchio, D., Boon, N., 2013. Environmental conditions and community evenness determine the outcome of biological invasion. Nat. Commun. 4 (1), 1383.

Dittmar, T., Koch, B., Hertkorn, N., Kattner, G., 2008. A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. Limnol. Oceanogr. Methods 6 (6), 230–235.

Dixon, P., 2003. VEGAN, a package of R functions for community ecology. J. Veg. Sci. 14 (6), 927–930.

Ercumen, A., Gruber, J.S., Colford, J.M., 2014. Water distribution system deficiencies and gastrointestinal illness: a systematic review and meta-analysis. Environ. Health Perspect. 122 (7), 651–660.

Favere, J., Waegenaar, F., Boon, N., De Gusseme, B., 2021. Online microbial monitoring of drinking water: how do different techniques respond to contaminations in practice? Water Res. 202, 117387.

Fontanazza, C.M., Notaro, V., Puleo, V., Nicolosi, P., Freni, G., 2015. Contaminant intrusion through leaks in water distribution system: experimental analysis. Procedia Eng. 119, 426–433.

Guimera, R., Nunes Amaral, L.A., 2005. Functional cartography of complex metabolic networks. Nature 433 (7028), 895–900.

Gunter, H., Bradley, C., Hannah, D.M., Manaseki-Holland, S., Stevens, R., Khamis, K., 2023. Advances in quantifying microbial contamination in potable water: potential of fluorescence-based sensor technology. WIREs Water. 10 (1), e1622.

Hammes, F., Berney, M., Wang, Y., Vital, M., Koster, O., Egli, T., 2008. Flow-cytometric total bacterial cell counts as a descriptive microbiological parameter for drinking water treatment processes. Water Res. 42 (1–2), 269–277.

Hammes, F.A., Egli, T., 2005. New method for assimilable organic carbon determination using flow-cytometric enumeration and a natural microbial consortium as inoculum. Environ. Sci. Technol. 39 (9), 3289–3294.

Haydar, S., Arshad, M., Aziz, J.A., 2009. Evaluation of drinking water quality in urban areas of Pakistan: a case study of Southern Lahore. Pakistan J. Eng. Appl. Sci. 5, 16–23.

Heibati, M., Stedmon, C.A., Stenroth, K., Rauch, S., Toljander, J., Save-Soderbergh, M., Murphy, K.R., 2017. Assessment of drinking water quality at the tap using fluorescence spectroscopy. Water Res. 125, 1–10.

Hou, C., Chen, L., Dong, Y., Yang, Y., Zhang, X., 2022. Unraveling dissolved organic matter in drinking water through integrated ozonation/ceramic membrane and biological activated carbon process using FT-ICR MS. Water Res. 222, 118881.

Hunter, P.R., Chalmers, R.M., Hughes, S., Syed, Q., 2005. Self-reported diarrhea in a control group: a strong association with reporting of low-pressure events in tap water. Clin. Infect. Dis. 40 (4), e32–e34.

Karim, M.R., Abbaszadegan, M., Lechevallier, M., 2003. Potential for pathogen intrusion during pressure transients. J. AWWA 95 (5), 134–146.

Kauppinen, A., Pitkanen, T., Al-Hello, H., Maunula, L., Hokajarvi, A.M., Rimhanen-Finne, R., Miettinen, I.T., 2019. Two drinking water outbreaks caused by wastewater intrusion including sapovirus in Finland. Int. J. Environ. Res. Public Health 16 (22), 4376.

Kellerman, A.M., Dittmar, T., Kothawala, D.N., Tranvik, L.J., 2014. Chemodiversity of dissolved organic matter in lakes driven by climate and hydrology. Nat. Commun. 5, 3804.

Kinnunen, M., Dechesne, A., Proctor, C., Hammes, F., Johnson, D., Quintela-Baluja, M., Graham, D., Daffonchio, D., Fodelianakis, S., Hahn, N., Boon, N., Smets, B.F., 2016. A conceptual framework for invasion in microbial communities. ISME J. 10 (12), 2773–2779.

Kwesiga, B., Pande, G., Ario, A.R., Tumwesigye, N.M., Matovu, J.K.B., Zhu, B.P., 2017. A prolonged, community-wide cholera outbreak associated with drinking water contaminated by sewage in Kasese District, western Uganda. BMC Public Health 18 (1), 30.

LeChevallier, M.W., Gullick, R.W., Karim, M.R., Friedman, M., Funk, J.E., 2003. The potential for health risks from intrusion of contaminants into the distribution system from pressure transients. J. Water Health 1 (1), 3–14.

Liu, G., Bakker, G.L., Li, S., Vreeburg, J.H., Verberk, J.Q., Medema, G.J., Liu, W.T., Van Dijk, J.C., 2014. Pyrosequencing reveals bacterial communities in unchlorinated drinking water distribution system: an integral study of bulk water, suspended solids, loose deposits, and pipe wall biofilm. Environ. Sci. Technol. 48 (10), 5467–5476.

Liu, S., Gunawan, C., Barraud, N., Rice, S.A., Harry, E.J., Amal, R., 2016. Understanding, monitoring, and controlling biofilm growth in drinking water distribution systems. Environ. Sci. Technol. 50 (17), 8954–8976.
Logue, J.B., Stedmon, C.A., Kellerman, A.M., Nielsen, N.J., Andersson, A.F., Laudon, H., Lindstrom, E.S., Kritzberg, E.S., 2016. Experimental insights into the importance of aquatic bacterial community composition to the degradation of dissolved organic matter. ISMe J. 10 (3), 533–545.
Luo, L.W., Wu, Y.H., Yu, T., Wang, Y.H., Chen, G.Q., Tong, X., Bai, Y., Xu, C., Wang, H.B., Ikuno, N., Hu, H.Y., 2021. Evaluating method and potential risks of chlorine-resistant bacteria (CRB): a review. Water Res. 188, 116474.
M. Aenab, A., K. Singh, S., 2012. Evaluation of drinking water pollution and health effects in Baghdad, Iraq. J. Environ. Prot. (Irvine, Calif) 03 (06), 533–537.
Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. 17 (1).
Mitch, W.A., Richardson, S.D., Zhang, X., Gonsior, M., 2023. High-molecular-weight by-products of chlorine disinfection. Nat. Water 1 (4), 336–347.
Nescerecka, A., Hammes, F., Juhna, T., 2016. A pipeline for developing and testing staining protocols for flow cytometry, demonstrated with SYBR Green I and propidium iodide viability staining. J. Microbiol. Methods 131, 172–180.
NRC, 2006. Drinking Water Distribution Systems: Assessing and Reducing Risks. The National Academies Press, Washington, DC.
Nygard, K., Wahl, E., Krogh, T., Tveit, O.A., Bohleng, E., Tverdal, A., Aavitsland, P., 2007. Breaks and maintenance work in the water distribution systems and gastrointestinal illness: a cohort study. Int. J. Epidemiol. 36 (4), 873–880.
Odhiambo, M., Viñas, V., Sokolova, E., Pettersson, T.J.R., 2023. Health risks due to intrusion into the drinking water distribution network: hydraulic modelling and quantitative microbial risk assessment. Environ. Sci. Water Res. Technol. 9 (6), 1701–1716.
Osterholz, H., Singer, G., Wemheuer, B., Daniel, R., Simon, M., Niggemann, J., Dittmar, T., 2016. Deciphering associations between dissolved organic molecules and bacterial communities in a pelagic marine system. ISMe J. 10 (7), 1717–1730.
Prest, E.I., Hammes, F., Kotzsch, S., van Loosdrecht, M.C., Vrouwenvelder, J.S., 2013. Monitoring microbiological changes in drinking water systems using a fast and reproducible flow cytometric method. Water Res. 47 (19), 7131–7142.
Prest, E.I., Hammes, F., van Loosdrecht, M.C., Vrouwenvelder, J.S., 2016. Biological stability of drinking water: controlling factors, methods, and challenges. Front. Microbiol. 7, 45.
Ramseier, M.K., Peter, A., Traber, J., von Gunten, U., 2011. Formation of assimilable organic carbon during oxidation of natural waters with ozone, chlorine dioxide, chlorine, permanganate, and ferrate. Water Res. 45 (5), 2002–2010.
Raza, S., Shin, H., Hur, H.G., Unno, T., 2022. Higher abundance of core antimicrobial resistant genes in effluent from wastewater treatment plants. Water Res. 208, 117882.
Riley, M.R., Gerba, C.P., Elimelech, M., 2011. Biological approaches for addressing the grand challenge of providing access to clean drinking water. J. Biol. Eng. 5 (1), 2.
Säve-Söderbergh, M., Bylund, J., Malm, A., Simonsson, M., Toljander, J., 2017. Gastrointestinal illness linked to incidents in drinking water distribution networks in Sweden. Water Res. 122, 503–511.
Semenza, J.C., 2020. Cascading risks of waterborne diseases from climate change. Nat. Immunol. 21 (5), 484–487.
Sorensen, J.P.R., Nayebare, J., Carr, A.F., Lyness, R., Campos, L.C., Ciric, L., Goodall, T., Kulabako, R., Curran, C.M.R., MacDonald, A.M., Owor, M., Read, D.S., Taylor, R.G., 2021. In-situ fluorescence spectroscopy is a more rapid and resilient indicator of faecal contamination risk in drinking water than faecal indicator organisms. Water Res. 206, 117734.
Sorensen, J.P.R., Vivanco, A., Ascott, M.J., Gooddy, D.C., Lapworth, D.J., Read, D.S., Rushworth, C.M., Bucknall, J., Herbert, K., Karapanos, I., Gumm, L.P., Taylor, R.G., 2018. Online fluorescence spectroscopy for the real-time evaluation of the microbial quality of drinking water. Water Res. 137, 301–309.
Sousi, M., Liu, G., Salinas-Rodriguez, S.G., Knezev, A., Blankert, B., Schippers, J.C., van der Meer, W., Kennedy, M.D., 2018. Further developing the bacterial growth potential method for ultra-pure drinking water produced by remineralization of reverse osmosis permeate. Water Res. 145, 687–696.
Stedmon, C.A., Seredyńska-Sobecka, B., Boe-Hansen, R., Le Tallec, N., Waul, C.K., Arvin, E., 2011. A potential approach for monitoring drinking water quality from groundwater systems using organic matter fluorescence as an early warning for contamination events. Water Res. 45 (18), 6030–6038.
Tanentzap, A.J., Fitch, A., Orland, C., Emilson, E.J.S., Yakimovich, K.M., Osterholz, H., Dittmar, T., 2019. Chemical and microbial diversity covary in fresh water to influence ecosystem functioning. Proc. Natl. Acad. Sci. U.S.A. 116 (49), 24689–24695.
Teunis, P.F., Xu, M., Fleming, K.K., Yang, J., Moe, C.L., Lechevallier, M.W., 2010. Enteric virus infection risk from intrusion of sewage into a drinking water distribution network. Environ. Sci. Technol. 44 (22), 8561–8566.
van Elsas, J.D., Chiurazzi, M., Mallon, C.A., Elhottovā, D., Krištůfek, V., Salles, J.F., 2012. Microbial diversity determines the invasion of soil by a bacterial pathogen. Proc. Natl. Acad. Sci. 109 (4), 1159–1164.
Vang, O.K., Corfitzen, C.B., Smith, C., Albrechtsen, H.J., 2014. Evaluation of ATP measurements to detect microbial ingress by wastewater and surface water in drinking water. Water Res. 64, 309–320.
Vila, J.C.C., Jones, M.L., Patel, M., Bell, T., Rosindell, J., 2019. Uncovering the rules of microbial community invasions. Nat. Ecol. Evol. 3 (8), 1162–1171.
Vinas, V., Sokolova, E., Malm, A., Bergstedt, O., Pettersson, T.J.R., 2022. Cross-connections in drinking water distribution networks: quantitative microbial risk assessment in combination with fault tree analysis and hydraulic modelling. Sci. Total. Environ. 831, 154874.
Wang, H., Proctor, C.R., Edwards, M.A., Pryor, M., Santo Domingo, J.W., Ryu, H., Camper, A.K., Olson, A., Pruden, A., 2014. Microbial community response to chlorine conversion in a chloraminated drinking water distribution system. Environ. Sci. Technol. 48 (18), 10624–10633.
Xu, X., Liu, Y., Liu, S., Li, J., Guo, G., Smith, K., 2019. Real-time detection of potable-reclaimed water pipe cross-connection events by conventional water quality sensors using machine learning methods. J. Environ. Manage. 238, 201–209.
Yang, J., Jiang, H., Liu, W., Huang, L., Huang, J., Wang, B., Dong, H., Chu, R.K., Tolic, N., 2020a. Potential utilization of terrestrially derived dissolved organic matter by aquatic microbial communities in saline lakes. ISME J. 14 (9), 2313–2324.
Yang, J., Jiang, H., Liu, W., Huang, L., Huang, J., Wang, B., Dong, H., Chu, R.K., Tolic, N., 2020b. Potential utilization of terrestrially derived dissolved organic matter by aquatic microbial communities in saline lakes. ISMe J. 14 (9), 2313–2324.
Yang, X., Shang, C., Lee, W., Westerhoff, P., Fan, C., 2008. Correlations between organic matter properties and DBP formation during chloramination. Water Res. 42 (8–9), 2329–2339.
Yao, M., Zhang, Y., Dai, Z., Ren, A., Fang, J., Li, X., van der Meer, W., Medema, G., Rose, J.B., Liu, G., 2023. Building water quality deterioration during water supply restoration after interruption: influences of premise plumbing configuration. Water Res. 241, 120149.
Zark, M., Dittmar, T., 2018. Universal molecular structures in natural dissolved organic matter. Nat. Commun. 9 (1), 3178.
Zhou, L., Wu, Y., Zhou, Y., Zhang, Y., Xu, H., Jang, K.S., Dolfing, J., Spencer, R.G.M., Jeppesen, E., 2024. Terrestrial dissolved organic matter inputs drive the temporal dynamics of riverine bacterial ecological networks and assembly processes. Water Res. 249, 120955.