Microbial Community Structure and Diversity in Drinking Water Supply, Distribution Systems as Well as Household Point of Use Site in Addis Ababa City, Ethiopia

Bayable Atanfu (✉ bayableatnafu@gmail.com)

Addis Ababa University <https://orcid.org/0000-0001-7613-9015>

Adey Desta

Department of Microbial, Cellular and Molecular Biology, Addis Ababa University

Fassil Assefa

Department of Microbial, Cellular and Molecular Biology, Addis Ababa University

Research

Keywords: Distribution systems, Drinking water, Microbial community, Residual chlorine, Household point of use

Posted Date: January 28th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-153906/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Understanding ecology of microbiomes in drinking water distribution systems is the most important notion in delivering safe drinking water. Despite cultivation-based methods routinely

employed in monitoring drinking water quality, cultivation of specific indicator organisms alone is not always guarantee for assuring safe drinking water delivery. The presence of complex microbiomes in drinking water distribution systems affects treatment effectiveness leading to poor quality water which as a result affects health of human and animals. Drinking water treatment and distribution systems harbor various microbiota despite efforts made in improving water infrastructures and several waterborne diseases become serious problems in the water industry, specially, in developing Countries. Intermittent water supply, long-time of water storage, low water pressure in distribution systems, storage tankers and pipes as well as contaminated source water are among many of the factors responsible for low drinking water quality which in turn affecting health of people. The aim of this study was to explore microbial diversity and structure in water samples collected from source water, treated water, reservoirs, and several household points of use locations (taps). High throughput Illumina sequencing technology was employed by targeting V4 region of 16S rRNA following Illumina protocol to analyze the community structure of bacteria.

Results

The core dominating taxa were Proteobacteria followed by Firmicutes, Bacteroidetes and Actinobacteria. Gamma proteobacteria were dominant among other Proteobacterial classes across all sampling points. Opportunistic bacterial genera such as *Pseudomonas*, *Legionella*, *Klebsiella*, *Escherichia*, *Actinobacteria*, as well as eukaryotic microbes like *Cryptosporidium*, *Hartmanella*, *Acanthamoeba*, *Aspergillus*, and *Candida* were also the abundant taxa found along

the distribution systems. The shift in microbial community structure from source to point of use

locations were influenced by factors such as residual free chlorine, intermittent water supply and

long-time storage at the household. The shift in microbial community structure from source to point of use locations were influenced by factors such as residual free chlorine, intermittent water supply and long-time storage at the household.

Conclusions

The complex microbiota which was present in different sample sites receiving treated water from the two treatment plants (Legedadi and Gefersa) starting from source water to household point of consumption across the distribution systems in Addis Ababa brings drinking water quality problem which further causes significant health problems to both human and animal health. Treatment ineffectiveness,

disinfection inefficiency, poor maintenance actions, leakage of sewage and other domestic wastes are few among many other factors responsible for degraded drinking water quality in this study putting health at high risk which, this, leads to morbidity and mortality. Findings of this research provide important and baseline information to understand the microbial profiles of drinking water along source water and distribution systems.

Background

Centralized drinking water treatment and distribution is the primary attribute of safe water supply; however, it is also among the biggest technological challenges as these systems must keep pace with rapid population growth and increased drinking water demand that comes with growing economic prosperity. Despite the advances in the development of urban water infrastructure, waterborne disease is still common in the cities of developing countries due to intermittent treated water supply which leads to excessive storage, low pressure events, poor integrity of distribution system piping, and contaminated source waters [1-4]. The Sustainable Development Goals (SDGs) are targeting these challenges and propose to achieve universal and equitable access to safe and affordable drinking water for all by 2030 [5]. Ethiopia is in the forefront of the countries in Sub-Saharan Africa with insufficient access to safe drinking water [6, 7]. Addis Ababa, the capital city, suffers from increasing human population, unregulated urban growth [8], poor waste management practices [9], intermittent water supply services [10, 11], insufficient operation and maintenance of treatment and distribution systems [12]; which are driving forces for drinking water contamination.

Although more attention is being directed to understand the microbial challenges experienced by Addis Ababa's drinking water, few studies use culture-independent methods. For example, several studies on the quality of drinking water in various regions across Ethiopia and in Addis Ababa used culture-dependent techniques focusing on specific indicator organisms [13-15]. Despite their convenience for routine monitoring, culture-based methods tend to underestimate the diversity and potential presence of both opportunistic and pathogenic microorganisms due to poor cultivability [16] or growth competition by non-pathogenic heterotrophic plate count (HPC) bacteria [17]. Next-generation sequencing platforms are useful for comprehensive assessment of drinking water microbiota and has begun to be employed in both resource-rich and constrained settings [18]. By employing next generation sequencing techniques, it is possible to detect and identify the presence and diversity of pathogenic microorganisms that cannot be achieved with classical methods [19].

This study aimed to elucidate the microbial community present, their diversity and community structure across the water distribution systems starting from source through treatment and along the distribution systems to storage and neighborhood tap for the two drinking water treatment plants serving Addis Ababa.

The specific objectives of this study are: i) studying the diversity of microorganisms available across the water supply systems; ii) studying dynamics and community structure of microorganism across the

water supply systems. Illumina sequencing was used with both 16S and 18S targets. The microbial diversity across the different point of distribution lines and household point of use locations in the City were comprehensively investigated.

Materials And Methods

Study Site Description and Sample Locations

The study was conducted in Addis Ababa City, the Capital of Ethiopia which lies at an elevation of 2,300 MASL. The City largely relies on surface and ground water as the main sources for drinking, domestic and industrial purposes. There are two surface water treatment plants, Legedadi and Gefersa, that provide 54% of the City's treated water (Fig. 1). Legedadi, established in 1970 GC (Gregorian calendar), is located 30 km from Addis Ababa on the North-East part. It is the largest drinking water treatment plant with a treatment capacity of 195,000 m³/day, which delivers about 47% of the daily distributed water supply for Addis Ababa. Gefersa was established in 1940 GC and is located 20 kms North-West part of the City with a treatment capacity of 30,000 m³ /day and providing only 7% of the city's distributed water. Legedadi and Gefersa treatment plants receive surface water from Dire and Gefersa dams, respectively. The treatment plants use conventional treatment that includes pre-chlorination, coagulation, flocculation, sedimentation, sand filtration and post-chlorination, with a goal of maintaining an average chlorine residual of 0.8mg/l in the distribution system.

For this study, a total of 38 samples were collected from several locations from source to tap along both the Legedadi (n=22) and Gefersa (n=16) systems (Table S1): Samples include source water entering into the treatment plant (LS and GS, n=1 each); finished drinking water before entering into distribution lines (LF and GF, n=1 each); réservoirs in the distribution system that stored treated water (LR, n=7 and GR, n=2); and household taps inside individual houses (LT; n=8 and GT; n=4) and storage tankers of individual houses (LS; n= 5 and GS; n=8). Samples were collected in July 2015 based on established sampling procedures [20]. All glass sample bottles were sterilized by autoclave (121⁰C, 15 min), then supplemented with sodium thiosulfate to quench residual chlorine in duplicate 2 L water samples per location. Prior to sample collection from each reservoir (R samples) and household standpipe (T and S samples), the water was allowed to run for 10 minutes in order to ensure representative sampling and to avoid sampling stagnant water

Microbial analysis of water samples

DNA extraction from water samples. Two liters of water per location were filtered through a polycarbonate membrane filters (0.22 µm pore size EMD Millipore™ GTTP02500) and frozen immediately at -20 ° C. All filtered samples were shipped frozen using ice packs via overnight courier service to the Environmental Biotechnology Laboratory at the University of Michigan, USA, and arrived frozen. Each membrane filter was cut into four equally-sized pieces using a sterile knife and placed into a single vial to facilitate DNA extraction. Total genomic DNA was extracted from each vial using a Power Soil DNA

Isolation Kit (MoBio Laboratories, Carlsbad, CA) following the manufacturer's protocol and instructions. The extracted DNA of each sample with total volume 50 µl was ready for high throughput sequencing process. The concentration and purity of the extracted DNA was determined using NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) (Table S2) and stored at -20 °C until sequencing.

16s rRNA and 18s rRNA gene PCR amplification and Sequencing

Extracted DNA was submitted to the Core Genome Sequencing unit at the University of Michigan, USA. The hypervariable V4 region of the 16S rRNA gene was amplified from each sample using the dual-indexing sequencing strategy with primers 515F (5'-GTGCCAGCMGCCGCGGTAA-3') and 806R (5'-GGACTACHVGGGTWTCTAAT-3') (23, 24). PCR amplification was done in a total volume of 20 uL reaction mixtures containing 2 uL 10x PCR Buffer II AccuPrime HiFi, 11.85 uL sterile water, 0.15 uL polymerase (AccuPrime), 5 uL primer set (4 uM) and 1 uL sample DNA. PCR amplification conditions were: 2 min at 95°C for 1 cycle, at 20 sec for 95°C (30 cycles), 15 sec at 55°C (30 cycle), 5 min at 72°C (30 cycles), and 10 min at 72°C for 1 cycle. Library preparation, normalization, quality checking, cluster generation and sequencing were conducted in the core Genome Sequencing Lab of the UM according to the protocol of Illumina Miseq platform.

For eukaryotic pathogens detection, 18S rRNA gene fragments were amplified using universal primers (forward and reverse) Euk-A7F (5'-AACCTGGTTGATCCTGCCAGT-3') and Euk-570R (5'-GCTATTGGAGCTGGAATTAC-3') which target the V1–V3 region of the 18S rRNA gene. PCR amplification was done in a total working volume of 770 uL containing 654.5 uL of sample library (containing Accuprime Tag, PCR water and 1X buffer II), 15.4 uL PhiX V3 as a control, and 100.1 uL of sample DNA. Library preparation, cluster generation and sequencing were carried out using the Illumina Miseq platform.

Bioinformatic and Statistical Analysis

The 16s rRNA reads generated from the Illumina Miseq analysis were filtered, clustered, taxonomically assigned and generally curated using MOTHUR (version 1.35.1) following the Miseq SOP (https://www.mothur.org/wiki/MiSeq_SOP) [21]. Sequencing and PCR errors were reduced by using the `make.contigs` command. Furthermore, sequences that did not align to the SILVA database were removed through sequence screening, and preculturing of rare sequences were merged into larger sequences [22]. The *chimera.vsearch* command was used to detect and remove chimeric sequences through Mothur [23, 24]. Classification of high-quality sequences into OTUs at 3% distance (97% similarity) were carried out based on the Ribosomal Database Project (RDP) Training set (Version 9) [25]. Clustering of OTUs at 3% dissimilarity was done using the *cluster.split* command. Rarefaction analysis, Chao1, Shannon diversity and evenness indices were computed using MOTHUR version 1.35.1 [21]. Non-metric multidimensional (n-MDS) analysis based on the Bray-Curtis dissimilarity index was generated using PAST software (version 3) for clustering of sampling sites based on microbial composition similarity. The relative

abundance bar graphs and Box-Whisker plot of phylum, class, and genus levels of the 16s rRNA sequence were plotted using IBM-SPSS (version 23) and XLSTAT tools. Heatmaps based on OTUs at the genus level were generated using ClustVis tool (web tool for visualizing clustering of multivariate data) [26]. The 18s rRNA gene sequences were submitted to MG_RAST web-based sequence analysis pipeline (<http://metagenomics.anl.gov/>) for qualitative analysis of the presence of public health-relevant parasitic and free-living eukaryotic organisms [27]. 18S reads were used for detection of the organisms only.

Results And Discussion

A total of 949,941, quality reads were processed and 22 phyla were retrieved from 729,385 assembled sequences in the samples taken from 38 sample locations. Sequence clustering created 604 unique bacterial OTUs and 1230 eukaryotic OTUs. The rarefaction curves showed that most of the sampling points failed to reach plateau, which suggested sampling frequency and sequencing depth should be improved to get enough reads (Fig. S1). The nonmetric MDS ordination plot showed marked differences on the similarity of bacterial composition based on the distribution of OTUs (Fig. 2).

Based on the analysis, samples having similar bacterial composition clustered together whereas those having different microbial assembly are dispersed apart as shown in Fig. 2a and b. Shannon-Weiner (H') and inverse Simpson (D^{-1}) diversity indices calculated for samples from Legedadi water sources indicate a bacterial diversity ranging between 0.6 – 1.7 and 1.5-14.3 respectively whereas Gefersa water sources had 0.7-1.7 and 0.7-40.3 of Shannon-Weiner (H') and inverse Simpson (D^{-1}) indices respectively. Untreated Source water sample locations of Gefersa treatment plants (GS), reservoirs (LR2, LR4 and GR1), end of pipe/tap locations (LT3 and GT1), and storage tanker locations (LHS4 and GHS7), all from both treatment plants had a lower Shannon-Weiner (H') values less than one (Table 1).

Samples taken from the water (LF) leaving out Legedadi treatment plant had higher diversity and low richness compared to source water (LS), but water taken from the five reservoirs showed increasing trend of bacterial richness along their distance from the treatment plant. This might be due to loss and/ or rapid decay of chlorine residuals along the reservoirs located in various degree of proximity from the treatment plant where chlorination is applied. Evidently, LR5 located in the farthest end of the reservoirs had higher diversity and richness. Among the end-of-pipe water samples, increasing trend of diversity and richness was observed with respect to the reservoirs they are sourced from. For example, LT1 and LT2, which are sourced from LR1, showed increasing bacterial diversity (1.5 to 1.7) and richness (104 to 376).

The water (GF) leaving out Gefersa treatment plant showed slight decrease in richness compared with the source (GS) water. Water samples from the two reservoirs (GR1 and GR2), showed a slight increase in their diversity and richness across their distance from the treatment plant. Samples from the four-taps (GT1- GT4) did not show consistent diversity/richness pattern along their distance from the treatment plant.

The samples from storage tankers of both treatment plants showed variable degrees of diversity and richness. This could be explained by their difference in tanker material, age of tanker usage, age of the water held and degree of cleaning frequency of the tanker. The increasing trend of bacterial diversity and richness in tap water samples of this study may be due to leakage of pipes, development of biofilms in the wall of pipes so that bacteria may get protection against residual disinfectants [28].

Inconsistent with this study, more diverse microbial taxa across metropolitan drinking water distribution systems at different geographic environment were detected and factors like source water, physicochemical parameters and treatment process affect microbiota diversity across distribution systems [29].

Members of phyla *Firmicutes* and *Proteobacteria* dominate reservoirs of the two treatment plant systems

A total of 22 phyla were retrieved from the assembled sequences from the water samples. The untreated raw water (LS) from Legedadi treatment plant were dominated by members of the phylum *Bacteroidetes* followed by *Actinobacteria* accounting for 88% and 87% of relative abundance of the total sequences respectively. Raw water from Gefersa treatment plants was dominated by members of the phylum *Proteobacteria* accounting for 65% of the total sequences, followed by *Actinobacteria* which was 25% of the total sequences (Fig. 3). On the other hand, LR2 reservoir receiving water from Legedadi treatment plant were found to harbor members of the phyla *Proteobacteria* and *Firmicutes* represented by 50% and 21% respectively. The remaining reservoirs which received water from this treatment plant also had higher dominant of *Proteobacteria* and *Firmicutes* with varying relative abundance as compared to other members of phyla. Similarly, finished water leaving the Gefersa treatment plant were found to harbor members of the phyla *Proteobacteria* and *Bacteroidetes* represented by 57% and 24% respectively. From reservoir sample locations (GR1 and GR2), *Proteobacteria* were found in high dominance with relative abundance of 83% and 61% respectively followed by *Firmicutes* which accounted for 13% and 34% of the total sequences, respectively. In line with this present study, in a study done by Li *et al.*, the most dominant bacterial taxa in raw water and subsequent sampling locations including finished water through drinking water distribution systems were *Proteobacteria* and *Firmicutes* followed by *Actinobacteria* and *Bacteroidetes* [30]. Reservoirs of drinking water distribution systems are naturally very complex environments supporting the growth of diverse microbial assemblages which can adapt to different hydraulic conditions and nutrient availability [31] which supported the presence of dominant phyla *Proteobacteria*, *Firmicutes* and *Bacteroidetes* in this present study. Moreover, there was a finding which reported the predominance of *Actinobacteria* and *Bacteroidetes* after disinfection of drinking water at various reservoir locations and in household storage tanker after disinfection [32].

Persistently highest abundance of *proteobacteria* was observed irrespective of distance from the source

The bacterial taxonomical composition in this study revealed that in both treatment plants, *Proteobacteria* was the most predominant bacterial phylum in all the sampling sites clustered according to their sampling locations (Fig. 4). The presence of *Proteobacteria* in high dominance in locations far from treatment plants in this present study might be due to the presence of low amount of residual

disinfectant, which, therefore favors re-growth, multiplication, and dominance of *Proteobacteria* and other microbial groups [33]. Moreover, the dominance of *proteobacteria* among other phyla in drinking water distribution systems was due to its ability to adapt at low nutrient availability conditions and having the capacity of biofilm formation at different pipe material surfaces [30, 34, 35].

Phylum *Proteobacteria* in both near and far sampling locations of Legedadi treatment plant had significantly higher relative abundance ($P < 0.01$ -Kruskal-Wallis test) of 86% and 84 % respectively followed by *Bacteroidetes* (15% relative abundance) in near sample locations and *Actinobacteria* (17% relative abundance) in far sample locations. This clearly indicated a shift of bacterial community structure from near to far sample locations. Similarly, phylum *Proteobacteria* with a relative abundance of 81% and 87% respectively in both near and far sampling locations was the most dominant phylum followed by *Firmicutes* with a relative abundance of 11% (near) and 7% (far) respectively in Gefersa treatment plant.

A study on metagenomic bacterial community from drinking water supply systems showed similar high dominance of *Proteobacteria*, followed by *Bacteroidetes* and *Actinobacteria* at various residential sample locations of the supply systems [1] and the dominance of these phyla in samples taken from raw water sources of a treatment plant, reaction tank, settling pond and clear water tank in another study were reported [36].

High abundance of class *Gamma proteobacteria* was found across sample locations along distance from sources of treatment plants

The dominant class in both near and far sample locations of Legedadi treatment plant were class γ -*proteobacteria* which accounted a relative abundance of 78 % and 88% respectively (Fig. 5). The second dominant class were β -*proteobacteria* accounted a relative abundance of 32 % and 28% from the five classes respectively in near and far sample locations. Similarly, in water samples sourced from Gefersa, class γ -*proteobacteria* with a relative abundance of 74% and 75% followed by β -*proteobacteria* with a relative abundance of 21% and 66% in near and far sample locations respectively. Like the findings of this present study, there were researches done at different location on microbial community structure in drinking water distribution systems and reported that δ followed by β -*proteobacteria* were the most dominant bacterial groups with increased relative abundance starting from treatment plant down to distribution systems [16, 29, 37].

Along the sampling locations which were far from Legedadi treatment plant, the abundance of α , γ and β -*proteobacteria* increased and similarly sampling locations far from Gefersa treatment plant, the relative abundance of γ -*proteobacteria* and β -*proteobacteria* increased and this indicated that these bacterial groups may be able to reside and proliferate at different locations of the drinking water distribution systems without significantly affected by disinfectants [38].

Phylum *Proteobacteria* were highly abundant in running tap and storage tanker locations

Phylum *Proteobacteria* were the most dominant phylum in samples received water from both Legedadi and Gefersa treatment plants. Significantly the highest relative abundance (95%) of phylum *Proteobacteria* ($P < 0.05$ -Kruskal-Wallis Test) in running tap water followed by 87% in storage tankers was observed where there was wide range relative abundance in running tap water samples as compared to storage tankers received water from Legedadi (Fig. S2). Phylum *Firmicutes* was the second dominant phylum (44%) in Legedadi household tap water sample locations. Similarly, the relative abundance of *Proteobacteria* in tap water and storage tanker sample location of Gefersa treatment plant was 98% and 92% respectively and *Firmicutes* (50%) followed by *Bacteroidetes* (42%) were the next abundant phylum in tap water samples whereas *Firmicutes* (12%) followed by *Actinobacteria* (5%) were the dominant phyla across storage tanker samples.

Several studies in consistent with this finding reported phylum *Proteobacteria*, *Bacteroides*, *Actinobacteria* and *Firmicutes* as the most dominant phylum in samples taken from tap water of urban drinking water supply systems [39, 40] and hence this present study finding was not an eye-opening in its high-level detection.

The presence of bacterial taxa in wide range of their respective relative abundance at each sample locations in this study was an indication of change in bacterial community from running tap and storage tanker locations throughout the distribution systems. Moreover, high dominance of bacterial taxa in tapes and storage tankers was probably due to resistance mechanisms against disinfectants and forming biofilms inside pipe surfaces and inside the wall of storage tankers as this fact was supported by scholars who did microbial community structure investigation in treated drinking water [41, 42]. The high dominance of bacterial taxa in running tap in this study compared to storage tanker sample locations might be changing water pressure through piping systems and general hydraulics which intern cause detachment of some cells from biofilms and resuspension of sediment associated bacteria. The presence of bacterial taxa to a lesser extent in storage tankers, on the other hand, in this study might be due to low water temperature and water stagnation which deters microbial metabolism and regrowth [43-45].

The presence of phylum *Actinobacteria* as the least proportion in treated tap water may be due to its vulnerability to disinfectant chlorine as this phylum is naturally dominant in raw and bulk water [46]. The dominance of *Proteobacteria* over taxa such as *Bacteroidetes*, *Firmicutes* and *Actinobacteria* in treated water samples collected from both running tap water and storage tanker in this study may be due to availability of nutrients under oxidative disinfection process which gives selective advantage for *Proteobacteria* having wide range of metabolic versatility and fast growth rate and also fundamental resilience to disinfection [32, 46]. The phyla *Proteobacteria*, *Bacteroidetes*, *Firmicutes* and *Actinobacteria* are the most common dominating bacterial groups in drinking water treatment systems [3, 37, 47].

Household tap and storage tanker locations were conducive environments for the dominance of *Gamma proteobacteria*

In running tap water and storage tankers samples sourced from Legedadi, class *γ-proteobacteria* was most predominant among five identified classes between tap and storage tanker samples with relative

abundance of 92% and 88% respectively whereas in storage and tap samples sourced from Gefersa, class *γ-proteobacteria* with relative abundance of 71% and 97% respectively was observed.

The second abundant class was *β-proteobacteria* which accounted 28% and 46% in tap water and storage tanker locations of Legedadi treatment plant and similarly 72 % and 53 % in tap and storage sample locations of Gefersa treatment plants respectively (Fig. 6).

Moreover, *δ-proteobacteria* were the other class detected in high relative abundance (74%) in the storage sample locations of Gefersa treatment plant. Across all the tap water and storage tanker sample locations, there were wide range in the relative abundance of all classes. On the other hand, *δ-proteobacteria* followed by, *β* and *γ-proteobacteria* were the most common and predominant bacterial groups usually found in residential water samples and chlorinated drinking water systems [1, 48]. There was significant variation in the relative abundance *γ-proteobacteria* among the four classes at each clustered sample location ($P < 0.05$ -Kruskal-Wallis Test).

The dominance of *γ-proteobacteria* shifted from the lowest to highest relative abundance across all clustered samples of tap water and storage tanker sample locations compared to other classes. Although *α-proteobacteria* has competitive advantage over other *Proteobacterial* groups by its virtual nature of existing under low nutrient availability and its ability to degrade complex organic compounds for its nutrient need in disinfected water [46], the predominance of *γ-proteobacteria* in this study would be attributed to its survival under suppressed environments in addition to its ability of tolerance to different treatment processes and disinfectant agents [49-51]. Moreover, the relative abundance of *γ-Proteobacteria* along the flow of the treatment process in another similar study showed that there was an increased pattern as compared to *α* and *β-Proteobacteria* classes which showed a decreasing relative abundance across the different steps of a treatment plant [30].

The presence of *β-proteobacteria* relatively with high percentage in potable water supplies may be the result of nutrient availability and due to formation of biofilms [52, 53] since this bacterial group is very sensitive to low nutrient availability and high residual disinfectant as compared to other *Proteobacterial* classes. Although the effect of residual chlorine results in a dynamic shift of bacterial community structure [28-30, 32], the findings of this study showed *γ-proteobacteria* had comparatively selective advantage over other classes in stored and piped water samples after chlorination. From this, it can be inferred that there could be different mechanisms like high tolerance to disinfection stress and recovery during treatment.

***Pseudomonas* was the most dominant genus across the distribution systems**

The relative abundances of all the identified genera from both Legedadi and Gefersa treatment plants are depicted in Fig. 7 and Fig S3. The relative abundance of *Pseudomonas* in source water (LS) of Legedadi treatment plant was 71% and showed a slight increase (77%) in finished water (LF). The trend fluctuated across the subsequent reservoirs from 55% in LR1 to 87% in LR5, with highest level (95%) in LR4. Drinking water is commonly inhabited by *Pseudomonas* species as this genus is characterized by high

prevalence, wide distribution in drinking water and its antibiotic resistance capacity [54]. Across tap water samples, *Pseudomonas* was the most dominant in LT5 (98%) followed by LT6 (96%) and water sample from the storage tanker, LS5, found to contain 95% relative abundance (Fig. 6). Like wise in Gefersa tretamnet plants, *Pseudomonas* was the also the dominnat genus which accounted 98% in tap water sample (LT5) located far from the treatment plant compared to low relative abundance (8%) of source water (GS) and 86% relative abundance from finished water sample (GF). Next to finished water sample, reservoir GR1 was found to contain 96% relative abundance of *Pseudomonas* and storage samples, GS5 and GS7 respectively contained 84 % and 77% relative abundance of *Pseudomonas*. In drinking water distribution systems and pipes, *Pseudomonas* and *Acinetobacter* commonly reside coated with thin films (biofilm) as long as organic matter is available there which these genera cause illness in young and immunocompromised people [55]. Moreover, distribution systems characterized by intermittent water supply, low water pressure, intrusions of pathogens, metals, other chemicals through breakage and infrastructure inefficiency and presence of organic matter which reduce residual chlorine as well as long time storage of water in the household storage tanker, all are important mechanisms of water contamination by opportunistic pathogens like *Pseudomonas* [4]. In agreement with this finding, high relative abundance of *Pseudomonas* in chlorinated drinking water distribution systems was also reported [56].

The genus *Pseudomonas* was repeatedly described as the leading biofilm former bacterial groups in treated drinking waters [57] which hence might be basic factor for its' presence at high relative abundance in drinking water distribution systems of this study. Moreover, this genus has the ability to attach in pipe surfaces and get protection from the action of disinfection by forming extra cellular polymeric substances (EPS) helping for cell-to-cell communication in addition to attachment to pipe surfaces [56, 58].

On the other hand, *Acinetobacter* was the second most abundant genus in Legedadi treatment plant samples, LS, with a relative abundance of 25%, Tap water (LT1), 87% and storage tanker sample (LS2) with a relative abundance of 93% and similarly a relative abundance of 92% was observed in source water (GS) of Gefersa treatment plant. Furthermore, tap water sample (GT4), storage tanker samples (GS2, GS3, GS6 and GS8) had a relative abundance of 93%, 99%, 61%, 63% and 55% of *Acinetobacter* respectively. *Acinetobacter* is the most common causative agent of hospital acquired infections particularly respiratory infections in susceptible individuals [59]. Similarly, it was reported that *Acinetobacter* was the most abundant in centralized drinking water treatment plants after disinfection by chlorine and several of its strains were found to survive without being affected significantly [60].

Moreover, *Escherichia* with 9%, 38%, 6% and 19% was observed in finished water (LF) and three reservoirs (LR1, LR2 and LR3) respectively. In agreement with this study, several opportunistic pathogenic genera were identified from chlorinated drinking tap water with *Escherichia* as predominant genera [61]. *Legionella* with 3% relative abundance was found in source water (LS), and very small percentage from finished water (LF), few of the reservoir (LR5), tap (LT2) and storage tanker (LS3). Interestingly, *Legionella* was found in the finished water (GF) only and not detected from the subsequent samples in Gefersa.

Although in very small percentage, detection of *Legionella* in distribution systems of this study might be associated with presence of host protozoa, invasion of *Legionella* from biofilms inside the pipe surface and low impact of chlorination [62].

The presence of these bacterial genera which are known to contain pathogenic species in drinking water treatment systems, storage tanks, public faucets and individual tapes were also reported [57, 63-65]. The diversity and dynamics of microbial communities in centralized water treatment plant showed the dominant of potential pathogenic bacterial groups in different sampling locations [36]. The presence of these opportunistic pathogens in this study may be due to rapid recovery after chlorination, development of resistance to disinfection, disinfection inefficiency, leakage after treatment and low disinfectant residual across subsequent sample locations of drinking water distribution system [66]. Moreover, the ability to survive under high temperature (up to 40⁰C) and adapt at low nutrient conditions of soil, fresh water and drinking water systems [67] were other factors which strengthen the presence of opportunistic pathogenic groups (*Pseudomonas*, *Acinetobacter* and *Legionella*) in considerable abundance.

Sporadic distribution of Eukaryote communities was observed across sampling locations

Diverse eukaryotes were detected from all sampling locations starting from source water up to running tap and storage tanker samples across the distribution networks. The most important genera found across different sample locations of the distribution system were *Hartmannella*, *Cryptosporidium*, *Cryptococcus*, *Acanthamoeba*, *Aspergillus* and *Candida* (Fig. 8). *Hartmannella* was found in 28% of all the samples sourced from Legedadi treatment plant and 44% of samples sourced from Gefersa treatment plant. Furthermore, *Cryptosporidium* was also found in 22% of the samples in Legedadi and 38 % of samples in Gefersa. Moreover, the genera *Acanthamoeba* was found in 56% of samples from Legedadi and 57% of samples sourced from Gefersa treatment plant. *Acanthamoeba*, the most prominent parasitic protozoa transmitted through contaminated water and *Cryptosporidium*, a causative agent of diarrheal disease in children in developing countries are the main problems in urban drinking water supply systems [68].

The opportunistic fungal pathogens including *Aspergillus* was found in all sampling locations except source water (LLS) and reservoir (LR2) of Legedadi treatment plant and also finished water (GF) and first reservoir (GR1) receiving water from Gefersa treatment plant. *Candida* accounted 67% and 50% in Legedadi and Gefersa sample locations respectively whereas *Cryptococcus* were observed in 50% and 38% of the sampling locations of Legedadi and Gefersa as well. Although many species of fungi under the genera *Penicillium* and *Aspergillus* produce mycotoxins in food and beverages, interestingly there are species of *Aspergillus* which produces aflatoxin in cold drinking water storage tanks [69]. The human pathogen *Candida* species was reported in drinking water and surface water [70] and similarly, sequencing results from fungi in drinking water and biofilms showed a higher abundance of *Cryptococcus* [71]. A phytoplankton microalgae genus *Cryptomonas* was also detected in 50% and 44% of the sampling locations which received water from Legedadi and Gefersa treatment plants respectively.

The unusual growth of genus *Cryptomonas*, a microalgae phytoplankton, plays for the presence of fishy odor in drinking water and hence contributes for low quality water [72].

Hartmanella and *Acanthamoeba* serve as a host for many pathogenic waterborne organisms such as *Legionella*, *Acinetobacter*, *Clostridium* and *Mycobacterium* [36, 73] and these host organisms are believed to found in surface water, waste water and drinking water systems in tropical regions where water temperature is 30 °C [74]. Although this study hasn't established the chlorine sensitivity test, the presence of *Hartmanella* in samples analyzed in this study showed its resistance to disinfection. Similar assessment of drinking water treatment plants showed very high residual chlorine concentration (0.8-3 mg/l) in produced water was insufficient for the removal of cysts of *Acanthamoeba* and other free-living amoebas [75].

Targeting 16S rRNA genes of specified regions (V4) with short read sequencing platforms cannot achieve the taxonomic resolution at species and strain level which is afforded by sequencing the entire genome. Likewise, identification of microbial community using 16S rRNA gene at genus level may mask important levels of inter-genus population differences and heterogeneity, which remain inaccessible to short-read 16S rRNA gene-based analysis. Furthermore, as limitation of this study, under sampling in this study clearly implied further frequent sampling need to be conducted to get sufficient sequence reads and to get bacterial populations in the study area.

Conclusions

To our knowledge this study is the first in investigating the microbial community structure and dynamics of drinking water distribution systems in Addis Ababa City using Illumina Miseq sequencing techniques. The microbial community structure, diversity and dynamics of water samples collected from source water entering to treatment plants, finished water inside treatment plants, reservoirs, and from various household point of use pipes across the distribution system was investigated. Highly diverse microbial communities were detected across all sampling locations. Moreover, the community structure shifted throughout the distribution systems. The phylum *Proteobacteria* followed by *Firmicutes*, *Bacteroidetes* and *Actinobacteria* was the highest predominant bacterial group found in all the sampling locations. Unlike several researches which reported the predominance of *Alpha proteobacteria* in treated drinking water systems, *Gamma proteobacteria* was the predominant class found followed by *Alpha* and *Beta proteobacteria*. Potentially opportunistic pathogens like *Pseudomonas*, *Legionella*, *Escherichia*, *Klebsiella*, *Proteus* and *Acinetobacter* were the dominant bacteria genera found in almost all the sampling locations. Free living Protozoan organisms like *Hartmanella* and *Acanthamoeba* which are hosts for opportunistic pathogens like *Legionella* and pathogenic ones like *Cryptosporidium* were also found in the samples analyzed.

Abbreviations

GC-Gregorian calendar, HPC-heterotrophic plate count, MASL-meter above sea level, MG-RAST-Metagenomic rapid annotations using subsystems technology, OUT, Operational taxonomic units, SDG-Sustainable development goals, SOP-standard operating procedures

Declarations

Acknowledgments

The authors thank Prof. Nancy G. Love (Environmental and Water Resource Engineering) at the University of Michigan, USA, for providing laboratory facilities, providing significant input and for she did rearrangement of the conceptual framework of the manuscript. We, the authors also thank Chia-Chen Wu for her assistance during DNA extraction. The authors also thank Addis Ababa Water and Sewerage Authority for providing logistic and personnel assistance during sample collection.

Authors' contribution

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Bayable Atanfu Kassa, and primary data analysis was done by Bayable Atanfu Kassa and Adey Feleke Desta. Bayable Atanfu Kassa and Adey Feleke Desta did data interpretation with and Bayable Atnafu Kassa rewrote it. The first draft of the manuscript was written by Bayable Atanfu Kassa with significant input from Adey Feleke Desta and Fassil Assefa. Fassil Assefa did proof reading of the manuscript. All authors commented on previous versions of the manuscript and are responsible for the content of this paper.

Funding

This study was financially supported by NSF supplemental grant for international collaboration through Nancy G. Love and Office of graduate programs at Addis Ababa University.

Data Availability

Data beyond that provided in the paper and supplemental information are available from the corresponding author upon reasonable request.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Saleem F, Mustafa A, Kori JA, Hussain MS, Kamran Azim M. (2018) Metagenomic Characterization of Bacterial Communities in Drinking Water Supply System of a Mega City. *MICROBIOLOGY OF AQUATIC SYSTEMS* 76(4):899-910. doi.org/10.1007/s00248-018-1192-2.
2. Prest E, Hammes F, Koetzsch S, van Loosdrecht M, Vrouwenvelder JS. (2013) Monitoring microbiological changes in drinking water systems using a fast and reproducible flow cytometric method. *Water research* 47(2013):7131-7142. doi.org/10.1016/j.watres.2013.07.051
3. Lautenschlager K, Hwang C, Ling F, Liu W-T, Boon N, Köster O (2014) Abundance and composition of indigenous bacterial communities in a multi-step biofiltration-based drinking water treatment plant. *Water research* 62(2014):40-52. doi.org/10.1016/j.watres.2014.05.035.
4. Kumpel E, Nelson KL (2016) Intermittent Water Supply: Prevalence, Practice, and Microbial Water Quality. *Environ. Sci. Technol*;50(2):542-553. DOI: 10.1021/acs.est.5b03973.
5. UNICEF, WHO (2017) Progress on drinking water, sanitation and hygiene: 2017 update and SDG baselines. Geneva: World Health Organization (WHO) and the United Nations Children's Fund (UNICEF), https://www.unicef.org/publications/index_96611.html; Licence: CC BY-NC-SA 3.0
6. UNICEF, WHO (2015) Progress on sanitation and drinking water-2015 update and Milleenium development goal assessment, Geneva: World Health Organization. https://www.unicef.org/publications/index_82419.html; ISBN: 978-92-4-150329-7
7. Gemechu Ameya, Olifan Zewdie, Abdulhakim Mussema, Adugna Amante and Birhanie Asmera (2018) Bacteriological quality of drinking water obtained from main sources, reservoirs and consumers tap in Arba Minch town, Southern Ethiopia. *African Journal of Microbiology Research* 12(24):567-573. DOI: 10.5897/AJMR2018.8871.
8. Bisrat Kifle Arsiso, Gizaw Mengistu Tsidu, Stoffberg GH, and Tsegaye Tadesse (2017) Climate change and population growth impacts on surface water supply and demand of Addis Ababa, Ethiopia. *Climate Risk Management* 18(2017):21-33. doi.org/10.1016/j.crm.2017.08.004.
9. Hamere Yohannis and Eyasu Elias (2017) Contamination of Rivers and Water Reservoirs in and Around Addis Ababa City and Actions to Combat It. *Environment Pollution and Climate Change* 1(2): 1-12.
10. Dawit Kidane Mekonnen (2015) The effect of distribution systems on household drinking water quality in Addis Ababa, Ethiopia, and Christchurch, New Zealand [Original]. New Zealand: University of Canterbury; Un published original Master's Thesis; <http://hdl.handle.net/10092/10692>.
11. Metadel Adane, Bezatu Mengistie, Girmay Medhin, Kloos H and Worku Mulat (2017) Piped water supply interruptions and acute diarrhea among under-five children in Addis Ababa slums, Ethiopia: A

- matched case-control study. *PloS one* 12(7):e0181516. <https://doi.org/10.1371/journal.pone.0181516>.
12. Hailu Worku (2017) Rethinking urban water management in Addis Ababa in the face of climate change: An urgent need to transform from traditional to sustainable system. *Environmental Quality Management* 27(1):103-19. DOI: 10.1002/tqem.21512.
 13. Solomon Abera, Zeyinudin A, Biruktwit Kebede, Amare Deribew, Solomon Ali S and Endalew Zemene (2011) Bacteriological analysis of drinking water sources. *Afr J Microbiol Res* 5(18): 2638-2641. [https:// DOI: 10.5897/AJMR11.218](https://doi.org/10.5897/AJMR11.218).
 14. Milkiyas Tabor, Mulugeta Kibret and Bayeh Abera (2011) Bacteriological and physicochemical quality of drinking water and hygiene-sanitation practices of the consumers in Bahir Dar city, Ethiopia. *Ethiop J Health Sci* 21(1): 19-26. <https://doi.org/10.4314/ejhs.v21i1.69040>.
 15. Bayeh Abera, Belay Bezabih and Derese Hailu (2017) Microbial Quality of Community Drinking Water supplies: A Ten year (2004-2014) analyses in West Amhara, Ethiopia 10(2017):22-26. <http://dx.doi.org/10.1016/j.swaqe.2016.06.001>.
 16. Vaz-Moreira I, Egas C, Nunes OC, Manaia CM (2013) Bacterial diversity from the source to the tap: a comparative study based on 16S rRNA gene-DGGE and culture-dependent methods. *FEMS microbiology ecology* 83(2013):361-374. DOI: 10.1111/1574-6941.12002.
 17. Leclerc H (2003) Relationships between common water bacteria and pathogens in drinking-water. In: Bartram J, Cotruvo J, Exner M, Fricker C, Glasmacher A, editors. *Heterotrophic Plate Counts and Drinking-water Safety; The Significance of HPCs for Water Quality and Human Health*. London SW1H 0QS, UK: IWA Publishing; https://www.who.int/water_sanitation_health/publications/hpc/en/.
 18. Tan B, Ng C, Nshimiyimana JP, Loh LL, Gin KYH, Thompson JR. Next-generation sequencing (NGS) for assessment of microbial water quality: current progress, challenges, and future opportunities. *Front. Microbiol.* 2015;6(1027). doi: 10.3389/fmicb.2015.01027.
 19. Ge Y, Yu S (2017) Next-generation sequencing for investigating the diversity of microorganisms and pathogenic bacteria in a water source. *Tropical Journal of Pharmaceutical Research* 16(7):1503-1512. <http://dx.doi.org/10.4314/tjpr.v16i7.7>.
 20. E.W. Rice, R.B. Baird, A.D. Eaton, editors (2017) *Standard methods for the examination of water and wastewater*: American Public Health Association. American Public Health Association (APHA), American Water Works Association (AWWA) and Water and Environment Federation (WEF) <https://www.awwa.org/Store/Product-Details/productId/65266295>. ISBN: 9780875532875
 21. Schloss PDW, Sarah L Ryabin, Thomas Hall, Justine R Hartmann, Martin Hollister, Emily B Lesniewski, Ryan A Oakley, Brian B Parks, Donovan H Robinson, Courtney J (2009) *Introducing*

- mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and environmental microbiology* 75(23):7537-7541. doi:10.1128/AEM.01541-09.
22. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P (2012) SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic acids research* 41(2012):590-596. doi:10.1093/nar/gks1219.
 23. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27(16):2194-2200; doi: 10.1093/bioinformatics/btr381.
 24. Rognes T, Flouri T, Nichols B, Quince C, Mahé F (2016) VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4:e2584. DOI 10.7717/peerj.2584.
 25. Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ (2009) The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic acids research* 37(2009):141-145. doi:10.1093/nar/gkn879.
 26. Metsalu T, Vilo J (2015) ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Research* 43(2015):566-570. doi: 10.1093/nar/gkv468.
 27. Wilke A, Gerlach W, Harrison T, Paczian T, Trimble WL, Meyer F (2015) A RESTful API for Accessing Microbial Community Data for MG-RAST. *PLOS Comp Biology* DOI: 10.1371/journal.pcbi.1004008.
 28. Revetta R, Gomez-Alvarez V, Gerke T, Santo Domingo J, Ashbolt N (2016) Changes in bacterial composition of biofilm in a metropolitan drinking water distribution system. *Journal of applied microbiology* 121(1):294-305. doi: 10.1111/jam.13150.
 29. Gomez-Alvarez V, Humrighouse BW, Revetta RP, Santo Domingo JW (2015) Bacterial composition in a metropolitan drinking water distribution system utilizing different source waters. *Journal of water and health* 13(1):140-151. doi: 10.2166/wh.2014.057.
 30. Li C, Ling F, Zhang M, Liu W-T, Li Y, Liu W (2017) Characterization of bacterial community dynamics in a full-scale drinking water treatment plant. *Journal of Environmental Sciences* 51(2017):21-30. <http://dx.doi.org/10.1016/j.jes.2016.05.042>.
 31. Lin H, Zhang S, Gong S, Zhang S, Yu X (2015) Characterization, Microbial Community Structure, and Pathogen Occurrence in Urban Faucet Biofilms in South China. *BioMed res int* 2015(Article ID 401672). <http://dx.doi.org/10.1155/2015/401672>.
 32. Becerra-Castro C, Macedo G, Silva AM, Manaia CM, Nunes OC (2016) Proteobacteria become predominant during regrowth after water disinfection. *Science of The Total Environment* 573 (2016):313-323. <https://doi.org/10.1016/j.scitotenv.2016.08.054>.

33. Qin K, Struewing I, Domingo J, Lytle D, lu J (2017) Opportunistic Pathogens and Microbial Communities and Their Associations with Sediment Physical Parameters in Drinking Water Storage Tank Sediments. *Pathogens* 6(54). doi:10.3390/pathogens6040054.
34. Zeng DN, Fan ZY, Chi L, Wang X, Qu WD, Quan ZX (2013) Analysis of the bacterial communities associated with different drinking water treatment processes. *World J Microbiol Biotechnol* 29(9):1573-1584. DOI 10.1007/s11274-013-1321-5.
35. Qu J, Zhang Q, Zhang N, Shen L, Liu P, editors (2015). *Microbial Community Diversity in Water and Sediment of an Eutrophic Lake during Harmful Algal Bloom Using MiSeq Illumina Technology*. *International Conference on Advances in Environment Research* 87(2015). DOI: 10.7763/IPCBE. 2015. V87. 12.
36. Lin W, Yu Z, Zhang H, Thompson IP (2014) Diversity and dynamics of microbial communities at each step of treatment plant for potable water generation. *Water research* 52(2014):218-230. <http://dx.doi.org/10.1016/j.watres.2013.10.071>.
37. Pinto AJ, Xi C, Raskin L (2012) Bacterial community structure in the drinking water microbiome is governed by filtration processes. *Environ. Sci. Technol* 46(16):8851-8859. dx.doi.org/10.1021/es302042t.
38. Buse HY, Lu J, Lu X, Mou X, Ashbolt NJ (2014) Microbial diversities (16S and 18S rRNA gene pyrosequencing) and environmental pathogens within drinking water biofilms grown on the common premise plumbing materials unplasticized polyvinylchloride and copper. *FEMS Microbiol. Ecol* 88(2014):280-295. DOI: 10.1111/1574-6941.12294.
39. Hwang C, Ling F, Andersen GL, LeChevallier MW, Liu W-T (2012) Microbial community dynamics of an urban drinking water distribution system subjected to phases of chloramination and chlorination treatments. *Applied and environmental microbiology* 78(22):7856-7865. doi:10.1128/AEM.01892-12.
40. Stuken A, Haverkamp THA, Dirven H, Gilfillan GD, Leithaug M, Lund V (2018) Microbial community composition of tap water and biofilms treated with or without Copper-Silver ionization. *Environ.Sci. Technol* 52(6):3354-3364. DOI: 10.1021/acs.est.7b05963.
41. Tian J, Lu J, Zhang Y, Li J-C, Sun L-C, Hu Z-L (2014) Microbial Community Structures and Dynamics in the O3/BAC Drinking Water Treatment Process. *Int. J. Environ. Res. Public Health* 11(6):6281-6290. doi:10.3390/ijerph110606281.
42. McCoy ST, VanBriesen JM (2014) Comparing spatial and temporal diversity of bacteria in a chlorinated drinking water distribution system. *Environmental Engineering Science* 31(1):32-41. DOI: 10.1089/ees.2013.0174.

43. El-Chakhtoura J, Saikaly PE, van Loosdrecht MCM, Vrouwenvelder JS (2018) Impact of Distribution and Network Flushing on the Drinking Water Microbiome. *Front. Microbiol* 9(2205). doi: 10.3389/fmicb.2018.02205.
44. Lavanya V, Ravichandran S(2013) Microbial contamination of drinking water at the source and household storage level in peri-urban area of southern Chennai and its implication on health, India. *J Public Health* 21(5):481-488. DOI 10.1007/s10389-013-0573-8.
45. Rožej A, Cydzik-Kwiatkowska A, Kowalska B, Kowalski D (2015) Structure and microbial diversity of biofilms on different pipe materials of a model drinking water distribution systems. *World J Microbiol Biotechnol* 31(2015):37-47. DOI 10.1007/s11274-014-1761-6.
46. Li Q, Yu S, Li L, Liu G, Gu Z, Liu M (2017) Microbial Communities Shaped by Treatment Processes in a Drinking Water Treatment Plant and Their Contribution and Threat to Drinking Water Safety. *Front. Microbiol* 8(2465). doi: 10.3389/fmicb.2017.02465.
47. Li X, Duan J, Xiao H, Li Y, Liu H, Guan F (2017) Analysis of bacterial community composition of corroded steel immersed in Sanya and Xiamen seawaters in China via method of Illumina MiSeq sequencing. *Front. Microbiol* 8(1737). doi: 10.3389/fmicb.2017.01737.
48. Liao X, Chen C, Wang Z, Chang C-H, Zhang X, Xie S (2015) Bacterial community change through drinking water treatment processes. *Int. J. Environ. Sci. Technol* 12(6):1867-1874. DOI 10.1007/s13762-014-0540-0.
49. Mi Z, Dai Y, Xie S, Chen C, Zhang X (2015) Impact of disinfection on drinking water biofilm bacterial community. *Journal of Environmental Sciences* 37(2015):200-205. <http://dx.doi.org/10.1016/j.jes.2015.04.008>.
50. Hong PY, Hwang C, Ling F, Andersen GL, LeChevallier MW, Liu WT(2010) Pyrosequencing analysis of bacterial biofilm communities in water meters of a drinking water distribution system. *Applied and environmental microbiology* 76(16):5631-5635. doi:10.1128/AEM.00281-10.
51. Kormas KA, Neofitou C, Pachiadaki M, Koufostathi E (2010) Changes of the bacterial assemblages throughout an urban drinking water distribution system. *Environ. Monit. Assess.* 165(1-4):27-38. DOI 10.1007/s10661-009-0924-7.
52. Williams MM, Domingo JWS, Meckes MC, Kelty CA, Rochon HS (2004) Phylogenetic diversity of drinking water bacteria in a distribution system simulator. *Journal of applied microbiology* 96(5):954-964. doi:10.1111/j.1365-2672.2004.02229.x.
53. Mathieu L, Bouteleux C, Fass S, Angel E, Block JC (2009) Reversible shift in the alpha, beta and gamma-proteobacteria populations of drinking water biofilms during discontinuous chlorination. *Water research* 43(14):3375-3386. doi:10.1016/j.watres.2009.05.005.

54. Vaz-Moreira I, Nunes OC, Manaia CM (2012) Diversity and antibiotic resistance in *Pseudomonas* spp. from drinking water. *Science of The Total Environment* 426 (2012) :366-374. doi: 10.1016/j.scitotenv.2012.03.046
55. Gray NF (2014) Pathogen Control in Drinking Water (Chapter Thirty). *Microbiology of waterborne diseases* (Second Edition). London: Academic Press 537-569. <https://doi.org/10.1016/B978-0-12-415846-7.00030-5>.
56. Douterelo I, Sharpe R, Boxall J (2014) Bacterial community dynamics during the early stages of biofilm formation in a chlorinated experimental drinking water distribution system: implications for drinking water discolouration. *Journal of applied microbiology* 117(1):286-301. doi:10.1111/jam.12516.
57. Vaz-Moreira I, Nunes OC, Manaia CM (2017) Ubiquitous and persistent Proteobacteria and other Gram-negative bacteria in drinking water. *Science of The Total Environment* 586 (2017):1141-1149. <http://dx.doi.org/10.1016/j.scitotenv.2017.02.104>.
58. Bédard E, Prévost M, Déziel E (2016) *Pseudomonas aeruginosa* in premise plumbing of large buildings. *MicrobiologyOpen* 5(6):937-956. DOI: 10.1002/mbo3.391.
59. Kelly JJ, Minalt N, Culotti A, Pryor M, Packman A (2014) Temporal Variations in the Abundance and Composition of Biofilm Communities Colonizing Drinking Water Distribution Pipes. *PloS one* 9(5):e98542. doi:10.1371/journal.pone.0098542.
60. Ma X, Vikram A, Casson L, Bibby K (2017) Centralized Drinking Water Treatment Operations Shape Bacterial and Fungal Community Structure. *Environmental science & technology* 51(13):7648-7657. doi: 10.1021/acs.est.7b00768.
61. Bishankha S, Joshi TP, Joshi DR, Bhatta DR (2013) Assessment of microbial quality of chlorinated drinking tap water and susceptibility of gram negative bacterial isolates towards chlorine. *Nepal Journal of Science and Technology* 13(1):173-178. <https://doi.org/10.3126/njst.v13i1.7456>.
62. Lesnik R, Brettar I, Höfle M (2015) *Legionella* species diversity and dynamics from surface reservoir to tap water: from cold adaptation to thermophily. *The Isme Journal* 10(5):1064-1080. doi: 10.1038/ismej.2015.
63. Mukherjee N, Bartelli D, Patra C, Chauhan BV, Dowd SE, Banerjee P (2016) Microbial diversity of source and point-of-use water in rural Haiti–A Pyrosequencing-based metagenomic Survey. *PloS one* 11(12):e0167353. doi:10.1371/journal.pone.0167353.
64. Lau HY, Ashbolt NJ (2009) The role of biofilms and protozoa in *Legionella* pathogenesis: implications for drinking water. *Journal of applied microbiology* 107(2):368-378. doi:10.1111/j.1365-2672.2009.04208.x.

65. Shaw JL, Monis P, Weyrich LS, Sawade E, Drikas M, Cooper AJ (2015) Using amplicon sequencing to characterize and monitor bacterial diversity in drinking water distribution systems. *Applied and environmental microbiology* 81(18):6463-6473. doi:10.1128/AEM.01297-15.
66. Chao Y, Mao Y, Wang Z, Zhang T (2015) Diversity and functions of bacterial community in drinking water biofilms revealed by high-throughput sequencing. *Scientific Reports* 5(10044). DOI: 10.1038/srep10044.
67. Penna VTC, Martins SAM, Mazzola PG (2002) Identification of bacteria in drinking and purified water during the monitoring of a typical water purification system. *BMC Public Health* 2(13). <https://doi.org/10.1186/1471-2458-2-13>.
68. Omarova A, Tussupova K, Berndtsson R, Kalishev M, Sharapatova K (2018) Protozoan parasites in drinking water: A System approach for improved water, sanitation and hygiene in developing countries. *International journal of environmental research and public health* 15(3): pii: E495. doi:10.3390/ijerph15030495.
69. Hageskal G, Lima N, Skaar I (2008) The study of fungi in drinking water. *mycological research* 113 (2):165-172. <https://doi.org/10.1016/j.mycres.2008.10.002>.
70. Babič MN, Gunde-Cimerman N, Vargha M, Tischner Z, Magyar D, Veríssimo C (2017) Fungal contaminants in drinking water regulation? a tale of ecology, exposure, purification and clinical relevance. *International journal of environmental research and public health* 14(6):636. doi:10.3390/ijerph14060636.
71. Raghupathi PK, Zupančič J, Brejnrod AD, Jacquiod S, Houf K, Burmølle M (2018) Microbiomes in Dishwashers: analysis of the microbial diversity and putative opportunistic pathogens in dishwasher biofilm communities. *Applied and environmental microbiology* 84 (5): AEM. 02755-17. DOI :10.1128/AEM.02755-17.
72. Lee J, Rai PK, Jeon YJ, Kim K-H, Kwon EE (2017) The role of algae and cyanobacteria in the production and release of odorants in water. *Environmental Pollution* 227:252-262. <https://doi.org/10.1016/j.envpol.2017.04.058>.
73. Valster RM, Wullings BA, van den Berg R, van der Kooij D (2011) Relationships between free-living protozoa, cultivable *Legionella* spp., and water quality characteristics in three drinking water supplies in the Caribbean. *Applied and environmental microbiology* 77(20):7321- 7328. doi:10.1128/AEM.05575-11.
74. Belila A, El-Chakhtoura J, Saikaly PE, van Loosdrecht MCM, Vrouwenvelder JS (2017) Eukaryotic community diversity and spatial variation during drinking water production (by seawater desalination) and distribution in a full-scale network. *Environmental Science: Water Research & Technology* 3(1):92-105. DOI: 10.1039/c6ew00265j.

75. Al-Herrawy AZ, Gad MA (2017) Assessment of Two Different Drinking Water Treatment Plants for the Removal of Free-living Amoebae, Egypt. Iranian Journal of Parasitology 12(3):413-422.
<http://ijpa.tums.ac.ir>.

Tables

Table 1. Community diversity and richness estimators of the 16S rRNA amplicon sequences of samples from the two treatment plants along the distribution lines and their reservoirs. Note that for each sample name, the first letter denotes the source of water (G= Gefersa and L = Legedadi), the second letter denotes whether it is the untreated source water (S), finished treated water (F), water from reservoir (R), water from end of pipe/tap (T) and water from storage tanker (S). The descriptions of sample points of each location and their distance from the respective treatment plants is found at Table S1)

Source	Site type	Sampling code	Shannon (H')	Inv. Simpson (D-1)	Chao I (lci, hci)	Distance from treatment plants
Legedadi	Treatment Plant	LS	1	2.4	1476 (1420, 1551)	0
		LF	1.5	5.1	109 (100, 135)	0
	Reservoir	LR1	1.2	10.4	92 (73, 155)	15.6
		LR2	0.6	1.5	121 (115, 140)	21.3
		LR3	1.6	5.4	168 (158, 195)	22
		LR4	0.6	5.7	152 (138, 186.3)	23.7
		LR5	1.3	7.7	220 (206, 252)	27.5
		LR6	1.4	7.3	90 (81, 117)	23
		LR7	1.5	6.8	187 (179, 210)	23.5
	Tap water	LT1	1.5	4.8	104 (95, 130)	16.5
		LT2	1.7	14.2	376 (360, 410)	18
		LT3	0.6	2.2	271 (261, 297)	22.6
		LT4	1.2	4.2	538 (521, 569)	23.2
		LT5	1.3	8.6	100 (92, 125)	24
		LT6	1.4	5.4	224 (208, 261)	24.9
		LT7	1.7	6.3	139 (119, 184)	24.5
		LT8	1.6	5.6	96 (91, 111)	18.5
	Storage tanker	LHS1	1.1	3.8	91 (82, 117)	19.8

Gefersa		LHS2	1.5	8	108 (87, 161)	20.5
		LHS4	0.7	3.1	118 (107, 152)	23
		LHS5	1.3	4.2	102 (80, 165)	23.7
		LH56	1.5	5.7	205 (200,243)	18.5
	Treatment Plant	GS	0.9	7.7	1584 (1540, 1644)	0
		GF	1	40.3	1390 (1353, 1444)	0
	Reservoir	GR1	0.9	5.7	161 (145, 199)	8
		GR2	1.5	17.8	295 (274, 336)	9
	Tap water	GT1	0.7	4.3	202 (185, 245)	8.5
		GT3	1.4	4.3	104 (95, 130)	12
		GT4	1.2	4.1	89 (73, 140)	14
		GT5	1.4	3.8	159(144, 196)	19.5
	Storage tanker	GHS1	1.5	4.4	108 (9, 146)	7
		GHS2	1.4	6.3	74 (64, 104)	9.5
		GHS3	1.1	5.3	130 (121, 153)	11
		GHS5	1.6	1.3	101 (95, 122)	12
		GHS6	1.1	18.2	181 (16, 238)	14
		GHS7	0.7	4.3	63.67 (57, 87)	15.6
		GHS8	1.3	5.1	165.8 (152, 198)	20

Figure 2

Non-metric multidimensional scaling (n-MDS) plot based on Brey-Curtis similarity index to ordinate the similarity of each sample locations of both Legedadi (a) and Gefersa (b) treatment plants on the bacterial taxa composition

Figure 3

Relative abundance of most dominant phyla across source (LS, GS), treated (finished) (LF, GF) water and réservoir locations (LR1-LR5; GR1-GR2) of both Legedadi (upper panel) and Gefersa (lower panel) treatment plants

Figure 4

Relative abundance of the four most dominant phyla in samples considered near and far from both Legedadi (upper panel) and Gefersa (lower panel) treatment Plants. Near sample locations which received water from Legedadi were LR1, LT1, LT6, and LS5) and were in the range of 15.6-19.8 Km whereas far sample locations in the range of 23.7-27.5 km from the treatment plant were LS1, LT3, LT4 and LR5. Near sample locations received from Gefersa treatment plant were: GS3, GR1, GT1 and GR2 and were in the range of 7km-9km, whereas Gefersa far sampling locations were GS7, GT3, GS6 and GS8 and these sampling locations were 14-20kms range from the treatment plant. Descriptions of the boxplot: the horizontal line is the median; the box below and above the median is the 25th and 75th quartiles; the whiskers below and above the boxes are the minimum and maximum values within the interquartile ranges

Figure 5

Relative abundance of classes of phylum Proteobacteria across sample locations considered near and far from the site of both Legedadi (upper panel) and Gefersa (lower panel) treatment plants. Descriptions of the boxplot: the horizontal line is the median; the box below and above the median is the 25th and 75th quartiles; the whiskers below and above the boxes are the minimum and maximum values within the interquartile ranges

Figure 6

Relative abundance of classes of Phylum Proteobacteria across tap water and storage tanker clustered sample locations of both Legedadi (upper panel) and Gefersa (lower panel) treatment Plants. Descriptions of the boxplot: the horizontal line is the median; the box below and above the median is the 25th and 75th quartiles; the whiskers below and above the boxes are the minimum and maximum values within the interquartile ranges; whereas the asterisks are outliers in the data.

a

b

Figure 7

Bar graph showing the relative abundance of pathogenic and opportunistic pathogenic Genera under Class γ -Proteobacteria at each sample locations of both Legedadi (A) and Gefersa (B) treatment plants

Figure 8

Histogram heatmap of 18s rRNA gene showing the presence of eukaryotic organisms across the samples of Legedadi and Gefersa treatment plants. The green colors are indication of the presence of organisms whereas the light lemon color is indication of the absence of organisms.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryfiguresandTables.docx](#)