Journal of Water & Health

IWA PUBLISHING

Journal of Water and Health Vol 22 No 3, 536 doi: 10.2166/wh.2024.290

Bacterial community assessment of drinking water and downstream distribution systems in highland localities of Ecuador

C. Alfonso Molina [a,b,*], Cristian Quiroz-Moreno [c], Pablo Jarrín-V. [d], Magdalena Díaz [b,e,f], Elizabeth Yugsi [g], Jorge Pérez-Galarza [h,i] and Lucy Baldeón-Rojas [h,i]

[a] Facultad de Medicina Veterinaria y Zootecnia, Universidad Central del Ecuador, Cdla. Universitaria y Gaspar de Carvajal s/n., 170521 Quito, Ecuador
[b] Instituto de Investigación en Zoonosis (CIZ), Universidad Central del Ecuador, Cdla. Universitaria y Gaspar de Carvajal s/n., 170521, Quito, Ecuador
[c] Department of Horticulture and Crop Science, Ohio State University, 2021 Coffey Road, Columbus OH 43210
[d] Laboratorio de Secuenciamiento de Ácidos Nucleicos, Dirección de Gestión de la Innovación, Instituto Nacional de Biodiversidad INABIO, Pje. Rumipamba N341 y Av. de los Shyris, Quito, Ecuador
[e] Facultad de Ingeniería Química, Universidad Central del Ecuador, Ritter s/n y Bolivia, 170521 Quito, Ecuador
[f] Institute of Integrative Systems Biology (I2SysBio), University of Valencia and Consejo Superior de Investigaciones Científicas (CSIC), Carrer del Catedràtic Agustín Escardino Benlloch, 46980 Paterna, Valencia, Spain
[g] Centro de Biotecnología 'Dr Daniel Alkalay Lowitt', Universidad Técnica Federico Santa María, General Bari 699, 2390136, Valparaíso, Chile
[h] Facultad de Ciencias Médicas, Universidad Central del Ecuador, Iquique N14-121 y Sodiro, Quito, Ecuador
[i] Instituto de Investigación en Biomedicina, Universidad Central del Ecuador, Capitán Giovanni Calles y Derby, Quito, Ecuador
*Corresponding author. E-mail: camolina@uce.edu.ec

CAM, 0000-0003-3972-8528; CQ-M, 0000-0002-9069-9147; PJ-V, 0000-0002-7431-0381; MD, 0000-0002-8472-5951; EY, 0000-0003-3092-2658; JP-G, 0000-0003-2742-3727; LB-R, 0000-0002-0447-0136

ABSTRACT

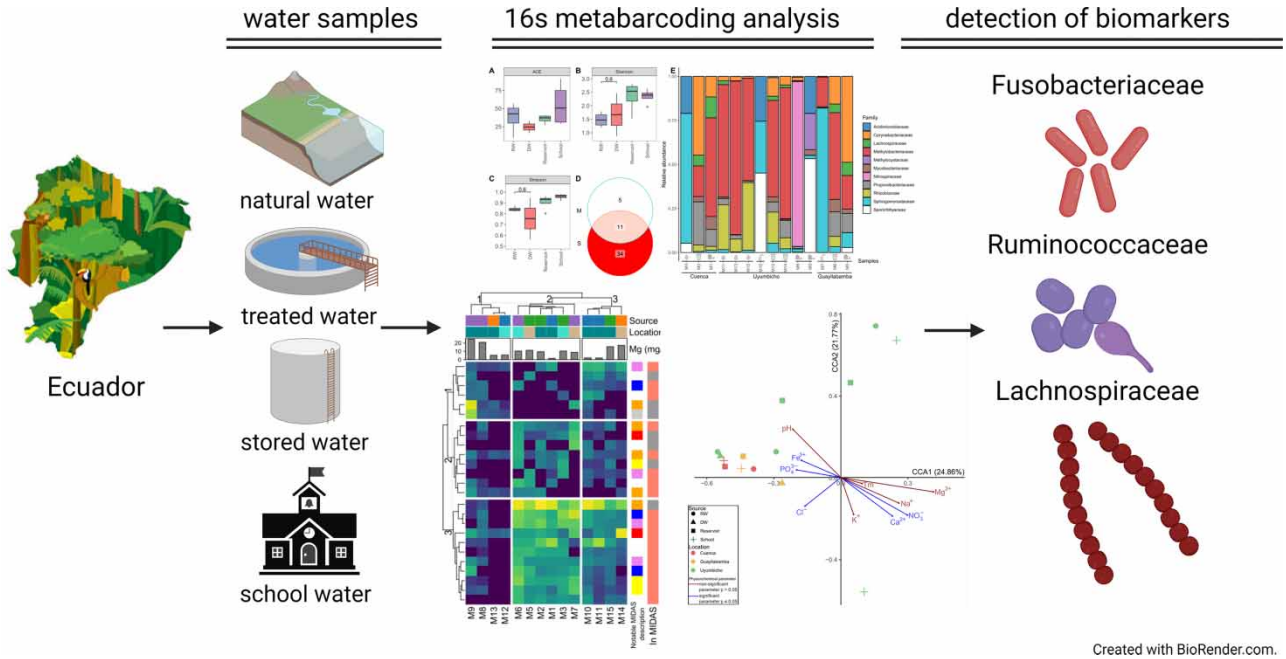
Bacterial communities in drinking water provide a gauge to measure quality and confer insights into public health. In contrast to urban systems, water treatment in rural areas is not adequately monitored and could become a health risk. We performed 16S rRNA amplicon sequencing to analyze the microbiome present in the water treatment plants at two rural communities, one city, and the downstream water for human consumption in schools and reservoirs in the Andean highlands of Ecuador. We tested the effect of water treatment on the diversity and composition of bacterial communities. A set of physicochemical variables in the sampled water was evaluated and correlated with the structure of the observed bacterial communities. Predominant bacteria in the analyzed communities belonged to Proteobacteria and Actinobacteria. The *Sphingobium* genus, a chlorine resistance group, was particularly abundant. Of health concern in drinking water reservoirs were Fusobacteriaceae, Lachnospiraceae, and Ruminococcaceae; these families are associated with human and poultry fecal contamination. We propose the latter families as relevant biomarkers for establishing local standards for the monitoring of potable water systems in highlands of Ecuador. Our assessment of bacterial community composition in water systems in the Ecuadorian highlands provides a technical background to inform management decisions.

Key words: bacterial communities, downstream distribution systems, drinking water, Lachnospiraceae, rural highlands, *Sphingobium*

HIGHLIGHTS

- This is the first report of 16S rRNA amplicon sequencing analysis of drinking water in rural communities in the highlands of Ecuador.
- On the basis of the presence and abundance of groups in the bacterial communities assessed in our study, we propose that the genera Fusobacteriaceae, Lachnospiraceae, and Ruminococcaceae can be considered as potential biomarkers for water quality monitoring programs with metabarcoding technologies.

GRAPHICAL ABSTRACT



1. INTRODUCTION

Access to drinking water is a fundamental human right; but as of 2015, the World Health Organization established that more than two billion people lack appropriate drinking water services (WHO 2017). According to the WHO, between 51 and 75% of the population in Ecuador has access to adequate drinking water services (WHO 2017). As a country with unique geographic features, such as the Andes and the Amazon basin, Ecuador has abundant freshwater resources and an extensive history of policies for the protection, use, and quality of freshwater (Hoogesteger *et al.* 2017; Joslin & Jepson 2018; Cisneros 2019). Since 2006, Ecuador has had a regulatory policy that outlines water quality standards, which was reviewed in 2011, 2014 (Kayser *et al.* 2015), and 2020 (Norma Técnica Ecuatoriana NTE INEN 11:08 2020). The definition of biological safety for drinking water in Ecuadorian policy is limited in scope. It considers the presence of three pathogens, fecal coliforms (<1 CFU/100 mL), *Cryptosporidium* spp. oocysts, and *Giardia* spp. cysts. Culture-dependent methods and microscopic evaluations are unavoidable requisites for detecting the three established biomarkers (APHA/AWWA/WEF2005). The Ecuadorian policy excludes a range of potential bacterial biomarkers that have been used to determine the source of human or poultry fecal contamination (FC), such as some of the species in Fusobacteriaceae, Lachnospiraceae, and Ruminococcaceae (Krause *et al.* 1999; Koskey *et al.* 2014).

We have followed the proposal by McLellan & Eren (2014) in considering those bacterial groups not formally used in national norms or standardized monitoring protocols as 'alternative indicators'; thus, our use of both terms, 'fecal coliforms' and 'indicators', refers to this particular argument. Fecal coliforms and indicators of fecal pollution are those organisms that naturally occur in the intestinal tracts of humans and other vertebrates, and their presence in water sources for human use is considered contamination and potentially detrimental to human health, usually as opportunistic pathogens (Bain *et al.* 2014; McLellan & Eren 2014). Fecal coliforms are the most widely employed microbial indicator to evaluate water quality and have historically led to the concept of public health protection (Rompré *et al.* 2002; Saxena *et al.* 2014). In Ecuador, coliform assessment has been used in the Galapagos islands to show the effect of centralized treatment systems for limiting contamination in drinking water (Gerhard *et al.* 2017). However, coliforms were reported as predominant in natural environments in Ecuador, for instance, exceeding human consumption limits in freshwater samples from the tropical Andes (Villa-Achupallas *et al.* 2018) and in natural water sources in the Amazon region and the northern Ecuadorian Pacific coast (Maurice *et al.* 2019). There are also reports that previously treated downstream drinking water supply systems could be contaminated with fecal coliforms (Levy *et al.* 2008). As a consequence of city wastewater discharge, high quantities of cefotaxime-resistant

Escherichia coli were reported along the watercourse of the Machángara River, which is the major urban river in Quito, the capital city of Ecuador (Ortega-Paredes *et al.* 2019).

Rural communities in Ecuador lack access to several rights and services, limitations enhanced by poverty (Mideros 2012). In 2008, the Constitution of Ecuador conferred on municipal governments the responsibility to provide water services to urban and rural settlements; 8 years later, about half of the country's rural population remained without drinking water services (Pozo *et al.* 2016). In rural areas, community water committees administer water provision services under the supervision of municipal governments and the technical support of the Health Minister provincial departments (Kayser *et al.* 2015). Besides, wastewater treatment is available only in a few main cities but is absent in rural areas (Maurice *et al.* 2019). Compared to urban systems, water distribution systems in rural Ecuador have financial and technical drawbacks that could influence water quality (Kayser *et al.* 2015). For instance, rural water systems in a cloud forest region showed a high concentration of fecal coliforms when compared to forested streams (Knee & Encalada 2013). A rural locality on the northern coast of Ecuador showed significantly high *E. coli* concentrations in surface and household drinking water (Levy *et al.* 2009). Therefore, poor water quality leads to health problems such as the high prevalence of intestinal parasites in children, particularly in rural areas where socioeconomic conditions are complex and challenging (Jacobsen *et al.* 2007). The Ecuadorian National Institute of Statistics (INEC) recently established that 36.7% of children under 5 years in Ecuador are exposed to *E. coli*-contaminated water sources, and cases in rural areas (56.5%) are almost double those in urban settings (25.9%) (INEC 2023).

By applying 16S rRNA metabarcoding, we determined the composition of the bacterial communities in drinking water treatment plants and the downstream supply systems in schools and reservoirs in Ecuador. Our samples included two rural communities and one urban settlement in Ecuador. We measured relevant bacterial community-shaping environmental variables as part of water physicochemical parameters. On the basis of their relevance to public health, we also evaluated the presence of potential biomarkers for FC across different water sources. Our results provide insights into microbiome composition assessments for informing decision-making on water management and public health strategies in Ecuador.

2. METHODS

2.1. Sample collection

We surveyed three locations, which were two rural water treatment plants and their associated distribution systems at Uyumbicho and Guayllabamba (Pichincha province) and one urban water treatment plant and its associated distribution system at the city of Cuenca (Azuay province). We collected samples from four sources: (1) the natural source before the water treatment plant (labeled as 'raw water' = RW), (2) drinking water outlets after the water treatment plant (labeled as 'DW'), (3) reservoir after the water treatment plant, and (4) tap water from elementary schools that were part of the distribution system (Figure 1). The transition between RW samples and DW samples marked the contrast between untreated and treated water samples. Fourteen water samples (1 L each) were collected from the three localities (Table 1) and stored in sterile bottles at 4 °C for further processing and analysis.

2.2. Physicochemical parameters

We selected 11 physicochemical water quality parameters to correlate with the observed structure of bacterial communities. These were determined according to the APHA/AWWA/WEF (2017), and they were calcium (Ca^{2+}), chloride (Cl^{-}), phosphates (PO_4^{3-}), total iron (Fe^{3+}), magnesium (Mg^{2+}), nitrates (NO_3^{-}), potassium (K^{+}), and sodium (Na^{+}). We measured all parameters in milligrams per liter. We determined pH and temperature (T_m) *in situ*, with a portable pH meter (Mettler-Toledo SevenGO, Millipore, USA).

2.3. DNA extraction and amplicon sequencing

We extracted metagenomic DNA with the DNeasy PowerWater kit (Cat. No. 14900-50 NF QIAGEN). Before amplicon sequencing, we stored isolated DNA at −80 °C. We quantified DNA quality and concentration with a Nanodrop 2000 (Thermo Scientific). We amplified a partial region of 500 bp that included the hypervariable regions V3 and V4 of the 16S rRNA gene, with an optimized primer pair and according to Klindworth's protocol (Klindworth *et al.* 2013). Subsequently, we sequenced the amplicon libraries of the 16S rRNA genes of the bacterial metagenome in 2 × 300 bp paired-end fragments. We used synthesis sequencing technology in the Illumina MiSeq platform (2000).

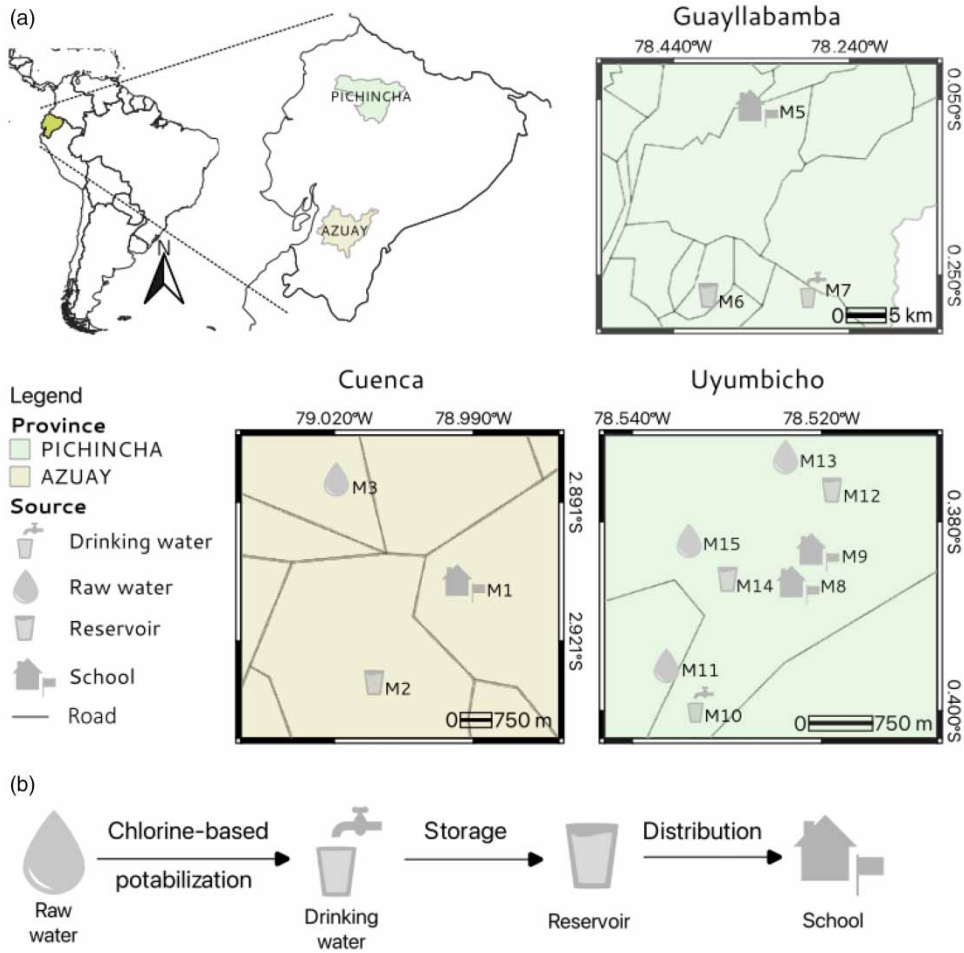


Figure 1 | (a) Sample locations, sample codification, and sample sources. (b) Sampling design along the drinking water distribution system from raw water sources to tap water in schools.

Table 1 | Samples and their locations

Sample	Province	Location	Source	Latitude	Longitude
M1	Azuay	Cuenca	School	−78.90	−2.91
M2	Azuay	Cuenca	Reservoir	−79.00	−2.93
M3	Azuay	Cuenca	RW	−79.00	−2.89
M5	Pichincha	Guayllabamba	School	−78.30	−0.06
M6	Pichincha	Guayllabamba	Reservoir	−78.40	−0.27
M7	Pichincha	Guayllabamba	DW	−78.20	−0.27
M8	Pichincha	Uyumbicho	School	−78.50	−0.39
M9	Pichincha	Uyumbicho	School	−78.50	−0.38
M10	Pichincha	Uyumbicho	DW	−78.50	−0.40
M11	Pichincha	Uyumbicho	RW	−78.50	−0.40
M12	Pichincha	Uyumbicho	Reservoir	−78.50	−0.38
M13	Pichincha	Uyumbicho	RW	−78.50	−0.37
M14	Pichincha	Uyumbicho	Reservoir	−78.50	−0.39
M15	Pichincha	Uyumbicho	RW	−78.50	−0.38

2.4. Amplicon data analysis

The sequence quality and trimming threshold for each sample were assessed in FastQC (Andrews 2010). We analyzed amplicon metagenomic libraries with the Divisive Amplicon Denoising Algorithm (DADA2) as it is implemented in the R package (Callahan *et al.* 2016a) and according to the methods proposed in the study by Callahan *et al.* (2016b), with modifications suitable for the data in the present study, as we explain later. We listed raw fastq files with the list.files() function, whereas the fastqPairedFilter() function was employed to remove low-quality reads. Truncation was applied at base pair positions 280 and 220 for forward and reverse reads, respectively. The 10 first nucleotides were trimmed at the start of both reads. We estimated a parametric error model through the learnErrors() function to remove independent errors within and between reads. We used the derepFastq() routine to dereplicate amplicon sequences. The discrimination of sequencing errors from real biological variation was performed with the dada() function and based on the previously estimated error model. We merged forward and reverse reads with the mergePairs() function. We obtained an analog of the operational taxonomic unit (OTU) table with the makeSequenceTable() function, consisting of a single row per OTU. We removed chimeras with the removeBimeraDenovo() function. We annotated the taxonomy with the assignTaxonomy() algorithm based on a Naive Bayesian classifier (Wang *et al.* 2007) and a dada2-formated Silva database. A second round of taxonomic annotation was performed with the dedicated amplicon metagenomics database for bacterial communities found in wastewater treatment plants (MIDAS 3) (Nierychlo *et al.* 2020). We processed the resulting community matrix in downstream analyses with the vegan, phyloseq, and ape packages in R (Oksanen *et al.* 2012; McMurdie & Holmes 2013; Paradis & Schliep 2019; R Core Team 2020).

We estimated alpha diversity indices (i.e., Abundance-based Coverage Estimator (ACE), Shannon, and Simpson) with the plot_richness() function in the vegan package. We applied a one-sided Wilcoxon rank-sum test on the contrast between raw and drinking water. Bacterial families were intersected for commonalities between raw and drinking water datasets, and the total number of common families between groups was estimated. We analyzed the structure of the community matrix through a canonical correspondence analysis (CCA) at the family level. We included the physicochemical parameters in the CCA as predictors. An iterative analysis of variance (ANOVA) test with 10,000 iterations evaluated the significance of physicochemical variables in the CCA model. The ANOVA test assessed the differences in the deviations of residuals in permutations of nested models (Legendre *et al.* 2011; Legendre & Legendre 2012). We estimated a neighbor joining tree for Lachnospiraceae and Ruminococcaceae since both families were of interest to the study (Callahan *et al.* 2016b; Paradis & Schliep 2019). We used packages 'muscle' (Edgar 2004) and 'phangorn' (Schliep 2011) for phylogenetic inference using neighbor joining-phylogram with 50 bootstrap repetitions. This phylogram explained the presence and abundance of particular genera and their relation to water sources and localities. Because of its superficial approach to robust phylogenetic inference, this tree should not be used for interpreting evolutive relationships. Absolute abundance values were $\log(x+1)$ transformed and represented in a heatmap with hierarchical clustering analysis for samples. The code corresponding to the applied methods is available at https://github.com/EcuadorianMP/Drinking-Water-Microbiome.

3. RESULTS

The raw sequence set is available as a NCBI bioproject with accession number PRJNA659797. Although diversity indexes process data patterns differently, a general tendency is possible to propose where bacterial diversity experienced a decrease after water treatment (i.e., the contrast between RW and DW) and then a noticeable increase at the reservoirs and schools (Figure 2(a)–2(c)). All Wilkinson nonparametric contrasts were significant between sources, except for the one between RW and DW for the Shannon and Simpson indexes (Figure 2(a)–2(c)). The Venn diagram depicts the unique and common bacterial families assigned to the obtained sequences, after the query results from the Silva and MiDAS databases (Figure 1(d)). The Silva database was six times more informative. There was a large variation in the relative abundance of the first 11 most abundant families across samples (Figure 2(e)). Observable patterns included a widespread presence and dominant abundance of Methylobacteriaceae across most of the samples, and Sphingomonadaceae, Sporichthyaceae, and Nitropiraceae were present in relatively large quantities on specific samples (M3, M7, M8, M9, and M10), which also included schools. Sharp differences were discernible in the transition across sample sources, particularly in the composition of families before and after water treatment or at the final stages of the downstream distribution system at schools.

The effect of water treatment was observable in changes in abundance across several bacterial groups. Sphingomonadaceae and Acidimicrobiaceae exhibited a reduction in abundance following water treatment in the Cuenca locality.

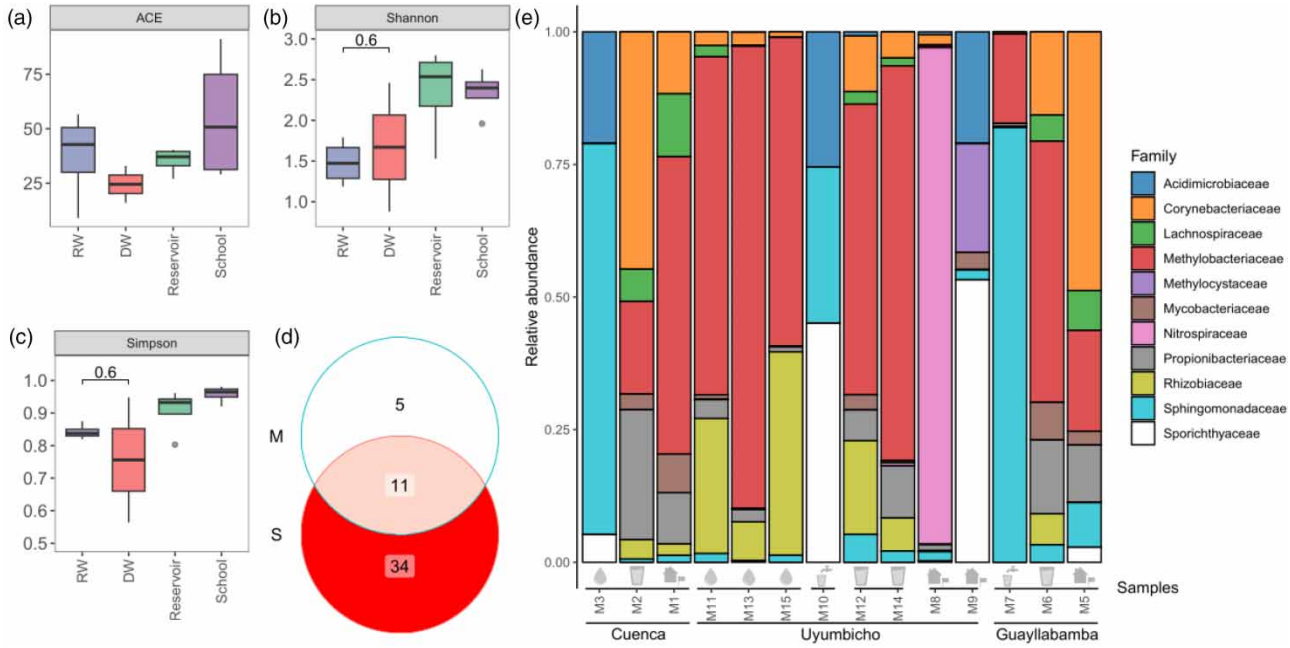


Figure 2 | Structure and diversity of bacterial communities across water samples before and after treatment. (a) ACE, (b) Shannon, and (c) Simpson alpha diversity indexes by sample. All Wilcoxon nonparametric contrast tests between samples were significant, except for the contrast between RW and DW for the Shannon and Simpson indexes ($p > 0.6$). (d) Venn diagram of the assigned families between the Silva-S and MiDAS-M databases. (e) Relative abundance by sample of the 11 most abundant families, with the taxonomic classification based on the Silva database.

Sphingomonadaceae was also noticeably reduced in abundance in the water reservoir at the Guayllabamba locality, contrasting with an increase in the abundance of Corynebacteriaceae in the school at the same locality. Methylobacteriaceae and Rhizobiaceae experienced lower abundances after water treatment at the Uyumbicho locality. Within the Sphingomonadaceae, genera *Sphingomonas* and *Novosphingobium* dominated in raw and treated water, respectively.

Based on the CCA results (Figure 3(a)), five variables showed statistical significance (p-value <0.05) for their correlation with the observed community structure, which were Mg ($p = 0.001$), Na ($p = 0.012$), temperature ($p = 0.007$), K ($p = 0.038$), and pH ($p = 0.003$). The locality of Uyumbicho showed a larger dispersion of samples when compared to Guayllabamba and Cuenca, meaning a more variable community structure of bacteria in the sampled water. This variability is also observable in Figure 2, where school samples for the Uyumbicho locality were dominated by both Sporichthyaceae and Nitrospiraceae (samples M8 and M9). As a vector in the spatial ordination of the CCA (Figure 3(a)), Cl^- is poorly correlated with the other physicochemical parameters and with the bacterial community structure observed in the samples. We assumed that Cl^- was part of the water treatment process, but we were not able to establish a clear pattern of its effects on bacterial community structure; however, the position of school and reservoir samples in the opposite gradient of concentration of this parameter was notable (Figure 3(a)). Most samples were in the opposite region of increasing concentrations for magnesium (Mg^{2+}), potassium (K^+), and sodium (Na^+) and higher temperature. Samples tended to cluster around a region with higher pH (Figure 3(a)).

Of particular interest to public health was the presence of the Lachnospiraceae and Ruminococcaceae families in the assessed water samples (Figure 3(b)). We inferred a phylogeny for representing the diversity of genera within both families. The phylogenetic tree of both families accounted for 27 OTUs at the genus level, although only 15 OTUs were classified into the genus level (Figure 3(b)). A total of 10 of the 15 genera were related to FC sources, of which 7 belonged to human sources and 3 to poultry-based FC. For example, *Lachnospira* is considered a biomarker of poultry FC and was found only in the Uyumbicho RW. In contrast, *Blautia*, a human FC biomarker, was found in almost all samples, except for those in Cuenca. *Roseburia* was abundant in schools and reservoirs across samples. *Faecalibacterium* was present across all sources, except for DW in Uyumbicho. It was particularly abundant in the school samples at Cuenca. *Ruminococcus* was abundant in

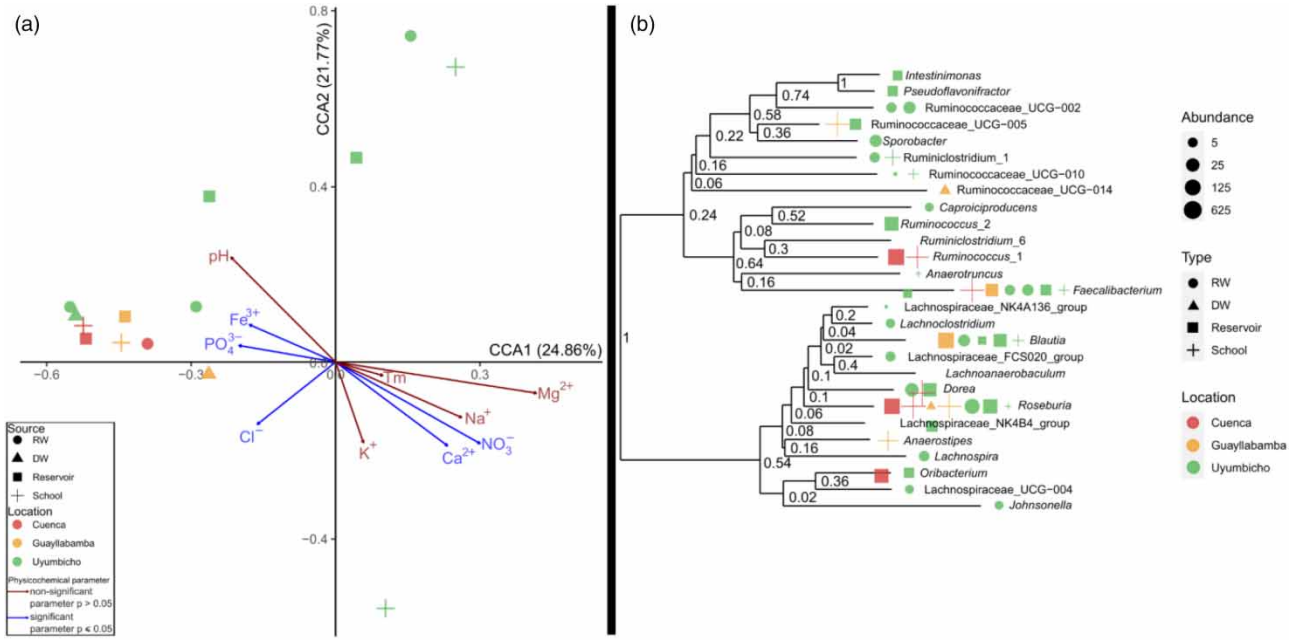


Figure 3 | (a) CCA of the bacterial communities at the family level. Environmental variables are linearly fit to the community structure (abundance and richness). According to the ANOVA test, Mg^{2+}, Na^{+}, K^{+}, and pH are statistically significant predictors ($p < 0.05$) of community structure. (b) A tree based on the neighbor joining algorithm for the Lachnospiraceae and Ruminococcaceae families, branch values are bootstrap support. The tree should not be used to interpret phylogenetic relationships (hence, the low branch bootstrap support), but to highlight the diversity of genera within these two health-relevant families in water treatment systems in Ecuador.

Cuenca and Uyumbicho, which was also present in samples from reservoir and school sources. *Anaerostipes* was notably abundant in Guayllabamba (Figure 3(b)).

Other families and genera detected for their abundance and significance to human health are worth mentioning, and we proceed to describe them next. Relative abundance values reported hereafter refer to the proportions within the selected subgroups listed in Figure 2(e). Mycobacteriaceae, which was present in all samples, had the highest relative abundance in water reservoirs at Cuenca (4%) and Guayllabamba (2%) (Figure 2). Mycobacteriaceae was also found in the Uyumbicho water reservoirs, with 0.4– 2% relative abundance. The MiDAS database provided species-level identifications for *Mycobacterium fallax* and *Mycobacterium insubricum*. Within Sphingomonadaceae (Figure 2), *Sphingomonas* was found in RW, schools, and reservoir samples, as the most abundant within Sphingomonadaceae and with a relative abundance range between 48 and 97% (percentages here represent proportions of the 11th most abundant families). In contrast, for DW samples, the dominant genus was *Novosphingobium*, which represented 89% of all genera in Sphingomonadaceae and 67% of the total relative abundance. Water treatment decreased the abundance of Fusobacteriaceae from 1% in RW to 0.02% in DW, indicating that the assessed treatment did not eradicate this bacterial family. The RW from Cuenca did not show the presence of Fusobacteriaceae; however, the sampled reservoir and school at Cuenca presented this family with a relative abundance of 4 and 1% respectively, which suggested contamination after potabilization.

As an overall graphical synthesis of the most salient patterns that resulted from the water samples survey, the heatmap shown in Figure 4 includes the 25 most abundant families across samples, their estimated abundance, relationship with the strongest and significant physicochemical parameter (Mg^{2+}), as determined by the CCA (Figure 3), and clustering patterns in terms of location, source, and notable description in the MIDAS database. Clustering patterns in the heatmap were not clear enough to determine consistent effects of the water treatment across either locations or sources.

4. DISCUSSION AND CONCLUSIONS

Through microbial metabarcoding, we have provided an assessment of the bacterial communities present in rural and urban water distribution systems in Ecuador, the effect of water potabilization, and the potential risks that may be present in drinking water. Although ours is the first assessment of its kind in Ecuador, we also acknowledge the limited scope of our study,

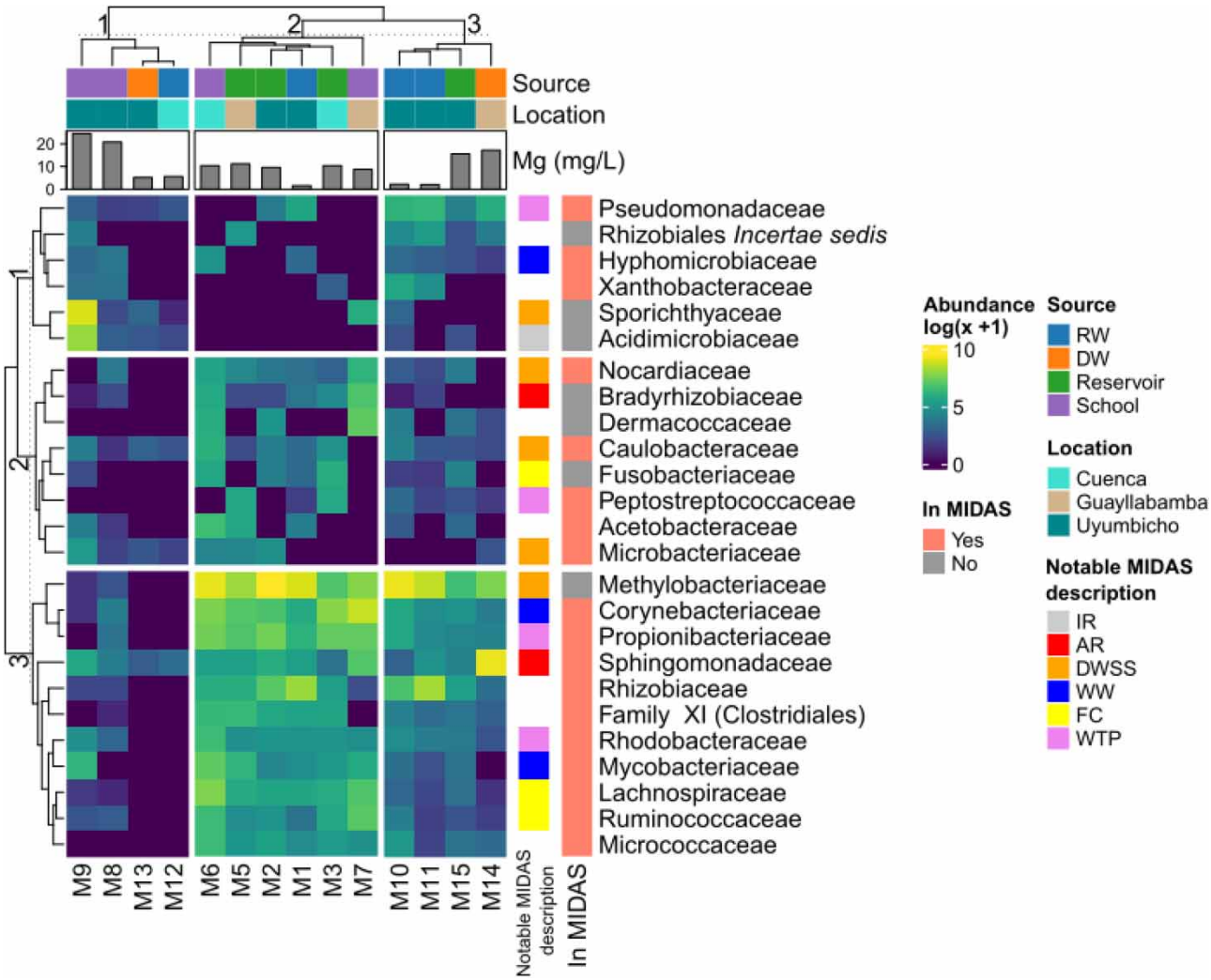


Figure 4 | Abundance heatmap of the first 25 most abundant families. The OTU's abundance was log-transformed and clustered according to their occurrence by the unweighted pair group method with arithmetic mean (UPGMA) hierarchical clustering. Annotations and color scales correspond to metabolic characteristics found in the literature and their presence in the MiDAS database.

particularly in terms of sample size and statistical coverage, which could have been reinforced by a longitudinal analysis in proper monitoring of water systems through seasons and administrative management intervals. We have discussed elsewhere the urgent and sensible challenges experienced in Ecuador for the study of microbiomes and microbiology in general (Díaz *et al.* 2021; Díaz *et al.* 2023). It is within this context that our assessment, although limited in its coverage, is important for contributing to strengthening the current state of microbial research in Ecuador, already limited and in urgent need of developing (Díaz *et al.* 2021), particularly in technological areas where human health is an issue.

The Ecuadorian National Standards Institute (known in Spanish as 'Servicio Ecuatoriano de Normalización INEN') maintains a technical norm for the microbiological examination of water samples, which was last updated in 2020 (Norma Técnica Ecuatoriana NTE INEN 11:08 2020). This technical norm focuses on the detection of fecal coliforms through culture methods, specifically of *Cryptosporidium* and *Giardia*. This standard in Ecuador, based on monitoring cultural fecal indicators and heterotrophic bacteria plate counts, is not different from current European standards, such as the European Directive 2020/2184 (Pinar-Méndez *et al.* 2022). In terms of bacterial community coverage, our study goes beyond the national and international standards and integrates with other initiatives elsewhere that push toward changing the culture plate paradigm in favor of more sensible and informative metabarcoding and massive parallel sequencing techniques (Li *et al.* 2018; Chang 2020). Yet, the use of high-throughput DNA sequencing technologies implies challenges as an effective

and accessible tool for monitoring the state of the drinking water. Particularly for developing countries, using metabarcoding monitoring includes budgetary and technological limitations for managing modern and dynamic nucleic acid sequencing tools and biomolecular information (Díaz *et al.* 2021). The applicability of 'a portable metagenomics toolbox', for monitoring drinking water quality, has been demonstrated as technically and economically viable for low-income countries (Acharya *et al.* 2020). Yet, Acharya *et al.* (2020) warn about the negative impact of self-imposed limitations in developing countries, from tariffs to freight costs on imported technology. Addressing these political and administrative issues is part of the necessary actions for establishing sound biomolecular monitoring programs in water treatment systems. Our study could also be considered as a pilot essay on the viability of modern and affordable approaches to sampling water quality in Ecuador. One advantage of the metabarcoding technique applied in our study is the detection of potential pathogens that are not recoverable in culture-dependent monitoring protocols, such as the many lineages in the Fusobacteriaceae, Lachnospiraceae, and Ruminococcaceae families.

One major challenge to applying 16S rRNA metabarcoding to monitor water quality, as part of a national standard, is the development of mechanisms that can confidently infer organismal abundance from DNA sequence abundance, effectively making the latter a consistent and reproducible quantitative application (Shelton *et al.* 2023). One meta-analysis reported a weak correlation between observed biomass and obtained number of sequences (Lamb *et al.* 2019). The current understanding among studies addressing the standardization of 16S rRNA metagenomic monitoring suggests that additional research is needed to establish a robust technological standard. This standard is anticipated to encompass comparisons against biomass estimates using traditional methods, evidence from multiple DNA markers, and the normalization of sequence abundance data with synthetic molecular standards (Zemb *et al.*, 2020; McElroy *et al.*, 2020; Ershova *et al.*, 2021; Skelton *et al.*, 2022).

We could not determine the role of Cl^- as it was not a significant physicochemical parameter and also did not correlate as expected with the bacterial community structure of the sampled water sources. Marked shifts in bacterial communities after chlorination have been observed in other settings (Hou *et al.* 2018), but the effect is not clearly defined as it can either favor or reduce bacterial diversity at various levels, promoting or inhibiting growth in specific organisms (Stanish *et al.* 2016; Dias *et al.* 2019). Our scope did not include a review of the chlorination strategies at the assessed water treatment plant; yet, the lack of statistical significance for Cl^- may warrant further assessments on the effectiveness of this water treatment component. Regarding other physicochemical parameters, a previous study found that bacterial community structure was positively correlated with turbidity, ammonia nitrogen, and total organic carbon (Li *et al.* 2017a, 2017b). One of the most comprehensive studies on correlations between bacterial community structure and physicochemical parameters identified Fe^{3+}, PO_4^{3-}, and NO_3^- as significant, which concurs with our study (Stanish *et al.* 2016). Nevertheless, each study applies a different scope in a distinct setting, and yet there is no universal guideline on which comparisons can be based across latitudes, cultural, and economical settings, and technological approaches to water treatment.

The effect of water treatment on bacterial community structure has been previously reported at other latitudes (Kormas *et al.* 2010; Kim *et al.* 2017; Zanacic *et al.* 2017; Zhang *et al.* 2017); however, it is the first time a study of this kind has been implemented for Ecuador. We have provided evidence that water treatment impacts the diversity of microorganisms, but does not necessarily eliminate those considered pathogens, which is a problem that has been previously reported in a controlled study in Ecuador (Levy *et al.* 2008). However, measurements of bacterial diversity do not always provide sufficient evidence to link potential pathogens in drinking water (Berry *et al.* 2006). Our findings reveal notable differences in the dominant bacterial phyla and genera compared to analyses conducted at other locations, particularly in the context of water treatment. Notably, a study on a drinking water treatment plant in Shanghai, China, reported increased taxonomic diversity in piped water, with Proteobacteria as abundant in microbial communities (Zeng *et al.* 2013). In contrast, Li *et al.* (2017a, 2017b) demonstrated a decrease in the proportion of bacteria during water treatment processes. A sequencing analysis of microbial communities in a drinking water treatment plant in Guangzhou, China, revealed the presence of opportunistic pathogens in finished water, consistent with our observations (Hou *et al.* 2018). Vavourakis *et al.* (2020) identified differences between microbial communities in drinking water distribution systems from rural and urban locations in the Netherlands, including seasonal effects. In a study conducted in Barcelona, Spain, Pinar-Méndez *et al.* (2022) analyzed the dynamics of a drinking water treatment plant, where Proteobacteria and Bacteroidota predominated in river water and throughout treatment. However, in the final drinking water, the selective pressure of treatment reduced diversity. Our results not only conform mostly to all previous findings on the effect of water treatment on reducing bacterial diversity but also confirm

that water treatment systems maintain potentially pathogenic bacteria. These findings reinforce the importance of vigilance in water treatment processes.

We have found evidence of human and animal FC, particularly in the families Fusobacteriaceae, Lachnospiraceae, and Ruminococcaceae (Figure 3). Fusobacteriaceae was detected in feces from pigs, horses, and cows (Koskey *et al.* 2014) and is a known inhabitant of the human gut, with association to disease (Robinson *et al.* 2020; Maynard 2023). The presence of fecal bacteria in water could be a consequence of upstream human settlements. The influence of poultry activity was previously reported in Ecuadorian water sources (Levy *et al.* 2012; Rao *et al.* 2015). Water quality in rural water systems was determined to be associated with their proximity to urban and agricultural activities (Knee & Encalada 2013). The detection of indicator bacteria in drinking water included the contribution of fecal matter in aquifers at a tropical high-altitude water treatment system in Ecuador (Espinosa *et al.* 2009). The persistence of certain bacterial families and genera associated with fecal sources, including those of human and poultry origin, even after water treatment, underscores potential contamination concerns. In summary, the identified bacterial families and genera, particularly those linked to FC, underscore potential risks to water safety.

The family Sphingomonadaceae, found in large abundances in some of the water samples, is of interest as it has been described in different water sources and has shown chlorine and antibiotic resistance (AR) (Sun *et al.* 2013; Narciso-da-Rocha *et al.* 2014). Sphingomonadaceae was found in drinking water treatment plants, tap water, and other water resources. Particularly, members of the genera *Sphingomonas*, *Sphingobium*, and *Novosphingobium* account for 60% of the identified genera (Vaz-Moreira *et al.* 2011). Notable AR was observed in *Sphingomonas* and *Sphingobium* (Vaz-Moreira *et al.* 2011). The Sphingomonadaceae family was also found in hospital tap water, with the genera *Sphingomonas*, *Novosphingobium*, and *Sphingobium* (Narciso-da-Rocha *et al.* 2014).

Lachnospiraceae and Ruminococcaceae are relevant biomarkers for the identification of FC (Krause *et al.* 1999; Koskey *et al.* 2014, Huws *et al.* 2011; Newton *et al.* 2011; Ponce-Terashima *et al.* 2014; Feng *et al.* 2018). At the genus level, human FC can be assessed by the presence of *Johnsonella* (Cotta *et al.* 2009), *Anaerostipes*, *Dorea* (Mclellan *et al.* 2013), *Caproiciproducens* (Chaplin *et al.* 2020), *Roseburia* (Wéry *et al.* 2010), *Blautia* (Koskey *et al.* 2014), and *Oribacterium* (Shen 2016). Moreover, ruminant FC can be evaluated by the presence of *Anaerotruncus* (Dowd *et al.* 2008) and *Faecalibacterium* (Shen *et al.* 2013). High concentrations of fecal bacteria were found in forests with minimal human impact, suggesting that they have a natural rather than anthropogenic source (Knee & Encalada 2013). Thus, the presence of Fusobacteriaceae, Lachnospiraceae, and Ruminococcaceae is not a definitive proof of environmental contamination with fecal material from human activity (Saxena *et al.* 2014). Not all bacteria in the families Lachnospiraceae and Ruminococcaceae can be inherently detrimental to human health. Some are part of the natural microbiome that contributes to homeostasis and health regulation. For example, *Roseburia intestinalis* produces butyrate in the human intestine and contributes to regulating various diseases (Nie *et al.* 2021). In terms of risk to human health, our findings are not definitive and should not be interpreted as a measure of risk factors. Risk classification to guarantee or prioritize interventions are commonly based on the number of indicator organisms, as direct counts, in a determined volume of sample (100 mL), which is a measurement that remains beyond the purpose and scope of our analysis (Bain *et al.* 2014).

The challenges of development, urbanization, and human population growth make the relationship between water quality and public policies unavoidable and intrinsically correlated. A national survey on water quality in Ecuador determined increased risks of contamination at the point of water consumption, with a significant deterioration of water quality between the years 2016 and 2019 nationwide, establishing that the most critical issue to address toward the national sustainable development goals for 2030 was water quality (Moreno *et al.* 2020). A focused analysis of the state of a major river in the Amazon region of Ecuador determined that it was unsuitable for human consumption, preservation of aquatic wildlife, and agricultural irrigation, particularly when considering national and international water quality standards (Villa-Achupallas *et al.* 2018). Public policy related to water security and quality in Ecuador was assessed by Kayser *et al.* (2015), through interviews with government officials, operators of water systems, and stakeholders they determined inadequate operation and services in many drinking water systems. The reasons identified for the deficiencies were related to insufficient technical capacity, inadequate water treatment oversight, and enforcement and lack of a robust financial system for service payments (Kayser *et al.* 2015).

The limitations in governance structure, bureaucratic hurdles, technical capacity, and financial resources to implement robust water quality surveillance must be addressed strategically and with priority. Public policies should address this problem by providing technical support to water committees and the associated communities living in rural areas. Our results

recommend implementing periodic water quality monitoring by analyzing the entire bacterial community and specific detection of potential pathogens using molecular methods. Rural water committees can readily adopt these procedures to enhance water management. Consequently, future initiatives should focus on fortifying monitoring infrastructure, promoting transparent communication, and ensuring the sustained viability of rural water committees. Our study provides a baseline on the microbial communities and occurrence of potential and opportunistic pathogens in two rural and one urban localities, with samples obtained along the water distribution systems. We hope this assessment could find value in the improvement of management guidelines and public health-related strategies that are relevant to water supply in Ecuador.

ACKNOWLEDGEMENTS

The authors extend their gratitude to the Corporación Ecuatoriana para el Desarrollo de la Investigación y Academia (CEDIA) and its CEPRA (November 2017) program for the financial support provided to the project 'Estudio del estado inmunológico de niños en edad escolar y su relación con el microbioma intestinal y con el agua potable que consumen'. We also thank Erika Rivadeneira for her help with sample collection. This work was an initiative of the Ecuadorian Microbiome Project (EcuMP).

AUTHOR CONTRIBUTIONS

C. A. M.: conceived the idea, experiment design, sample collection, laboratory experiments, and writing – review and editing. C. Q.-M.: bioinformatic and data analysis, and writing – review and editing. P. J.-V.: data analysis and writing – review and editing. M. D.: sample collection and physicochemical analysis. E. Y.: sample collection and laboratory experiments. J. P.-G.: experiment design and editing. L. B.-R.: conceived the idea, experiment design, writing – review and editing, and grant administration. All authors contributed to the article and approved the submitted version.

DATA AVAILABILITY STATEMENT

The code corresponding to the applied methods is available at https://github.com/EcuadorianMP/Drinking-Water-Microbiome. The raw FASTA sequence set is available as a NCBI bioproject with accession number PRJNA659797.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

Acharya, K., Blackburn, A., Mohammed, J., Haile, A. T., Hiruy, A. M. & Werner, D. 2020 Metagenomic water quality monitoring with a portable laboratory. *Water Research* **184**, 116112. https://doi.org/10.1016/j.watres.2020.116112.

Andrews, S. 2010 FastQC: A quality control tool for high throughput sequence data. Babraham Bioinformatics, Babraham, UK. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (accessed 1 July 2023).

APHA/AWWA/WEF. 2005 *Standard Methods for the Examination of Water and Wastewater*, 21st edn. American Public Health Association/American Water Works Association/Water Environment Federation, Washington, DC.

APHA/AWWA/WEF. 2017 *Standard Methods for the Examination of Water and Wastewater*, 23rd edn. American Public Health Association/American Water Works Association/Water Environment Federation, Washington, DC.

Bain, R., Cronk, R., Wright, J., Yang, H., Slaymaker, T. & Bartram, J. 2014 Fecal contamination of drinking-water in low- and middle-income countries: A systematic review and meta-analysis. *PLoS Medicine* **11** (5), e1001644. https://doi.org/10.1371/journal.pmed.1001644.

Berry, D., Xi, C. & Raskin, L. 2006 Microbial ecology of drinking water distribution systems. *Current Opinion in Biotechnology* **17**, 297–302. https://doi.org/10.1016/j.copbio.2006.05.007.

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A. & Holmes, S. P. 2016a DADA2: High-resolution sample inference from illumina amplicon data. *Nature Methods* **13**, 581.

Callahan, B. J., Sankaran, K., Fukuyama, J. A., McMurdie, P. J. & Holmes, S. P. 2016b Bioconductor workflow for microbiome data analysis: From raw reads to community analyses. *F1000Research* **5**, 1492. https://doi.org/10.12688/f1000research.8986.2.

Chang, J. 2020 Monitoring biodiversity and water pollution via high-throughput eDNA metabarcoding. *Berkeley Scientific Journal* **24**, 50–58. doi:10.5070/BS3242049349.

Chaplin, A. V., Sokolova, S. R., Shcherbakova, V. A., Suzina, N. E., Kochetkova, T. O., Goltsov, A. Y., Trofimov, D. Y. & Efimov, B. A. 2020 Hydrogeniiclostidium mannosilyticum gen. nov., sp. nov. isolated from human faeces. *International Journal of Systematic and Evolutionary Microbiology* **70** (2), 1210–1216. https://doi.org/10.1099/ijsem.0.003900.

Cisneros, P. 2019 What makes collaborative water governance partnerships resilient to policy change? A comparative study of two cases in Ecuador. *Ecology and Society* **24**, 1–29. https://doi.org/10.5751/ES-10667-240129.

Cotta, M. A., Whitehead, T. R., Falsen, E., Moore, E. & Lawson, P. A. 2009 *Robinsoniella peoriensis* gen. nov., sp. nov., isolated from a swine-manure storage pit and a human clinical source. *International Journal of Systematic and Evolutionary Microbiology* **59**, 150–155. https://doi.org/10.1099/ijs.0.65676-0.

Dias, V. C. F., Durand, A. A., Constant, P., Prévost, M. & Bédard, E. 2019 Identification of factors affecting bacterial abundance and community structures in a full-scale chlorinated drinking water distribution system. *Water* **11**, 627. https://doi.org/10.3390/w11030627.

Díaz, M., Jarrín, V. P., Simarro, R., Castillejo, P., Tenea, G. & Molina, C. A. 2021 The Ecuadorian microbiome project: A plea to strengthen microbial genomic research. *Neotropical Biodiversity* **7**, 223–237. https://doi.org/10.1080/23766808.2021.1938900.

Díaz, M., Monfort-Lanzas, P., Quiroz-Moreno, C., Rivadeneira, E., Castillejo, P., Arnau, V., Díaz, W., Agathos, S. N., Sangari, F. J., Jarrín, V. P. & Molina, C. A. 2023 The microbiome of the ice-capped Cayambe Volcanic Complex in Ecuador. *Frontiers in Microbiology* **14**, 1154815. https://doi.org/10.3389/fmicb.2023.1154815.

Dowd, S. E., Callaway, T. R., Wolcott, R. D., Sun, Y., McKeehan, T., Hagevoort, R. G. & Edrington, T. S. 2008 Evaluation of the bacterial diversity in the feces of cattle using 16S rDNA bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP). *BMC Microbiology* **8**, 1–8. https://doi.org/10.1186/1471-2180-8-125.

Edgar, R. C. 2004 MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**, 1792–1797. https://doi.org/10.1093/nar/gkh340

Ershova, E., Wangensteen, O., Descoteaux, R., Barth-Jensen, C. & Præbel, K. 2021 Metabarcoding as a quantitative tool for estimating biodiversity and relative biomass of marine zooplankton. *ICES Journal of Marine Science* **78**, 3342–3355. https://doi.org/10.1093/icesjms/fsab171.

Espinosa, A. C., Arias, C. F., Sanchez-Colán, S. & Mazari-Hiriart, M. 2009 Comparative study of enteric viruses, coliphages and indicator bacteria for evaluating water quality in a tropical high-altitude system. *Environmental Health* **8**, 1–10. https://doi.org/10.1186/1476-069X-8-49.

Feng, S., Bootsma, M. & McLellan, S. L. 2018 Human-associated Lachnospiraceae genetic markers improve detection of fecal pollution sources in urban waters. *Applied and Environmental Microbiology* **84**, 1–14. https://doi.org/10.1128/AEM.00309-18.

Gerhard, W. A., Choi, W. S., Houck, K. M. & Stewart, J. R. 2017 Water quality at points-of-use in the Galapagos islands. *International Journal of Hygiene and Environmental Health* **220**, 485–493. https://doi.org/10.1016/j.ijheh.2017.01.010.

Hoogesteger, J., Tiaguaro-Rea, Y., Rap, E. & Hidalgo, J. 2017 Scalar politics in sectoral reforms: Negotiating the implementation of water policies in Ecuador (1990–2008). *World Development* **98**, 300–309. https://doi.org/10.1016/j.worlddev.2017.04.036.

Hou, L., Zhou, Q., Wu, Q., Gu, Q., Sun, M. & Zhang, J. 2018 Spatiotemporal changes in bacterial community and microbial activity in a full-scale drinking water treatment plant. *Science of the Total Environment* **625**, 449–459. https://doi.10.1016/j.scitotenv.2017.12.301.

Huws, S. A., Kim, E. J., Lee, M. R., Scott, M. B., Tweed, J. K., Pinloche, E., Wallace, R. J. & Scollan, N. D. 2011 As yet uncultured bacteria phylogenetically classified as *Prevotella*, Lachnospiraceae *incertae sedis* and unclassified Bacteroidales, Clostridiales and Ruminococcaceae may play a predominant role in ruminal biohydrogenation. *Environmental Microbiology* **13** (6), 1500–1512. https://doi.org/10.1111/j.1462-2920.2011.02452.x.

INEC. 2023 Boletín Técnico Nro. 01-2023-ENDI, Documento Metodológico de la Encuesta Nacional sobre Desnutrición Infantil 2022-2023, Quito, Ecuador.

Jacobsen, K. H., Ribeiro, P. S., Quist, B. K. & Rydbeck, B. V. 2007 Prevalence of intestinal parasites in young Quichua children in the highlands of rural Ecuador. *Journal of Health, Population and Nutrition* **25**, 399–405. https://doi.org/10.3329/jhpn.v25i4.618.

Joslin, A. & Jepson, W. 2018 Territory and authority of water fund payments for ecosystem services in Ecuador's Andes. *Geoforum* **91**, 10–20. https://10.1016/j.geoforum.2018.02.016.

Kayser, G. L., Amjad, U., Dalcanale, F., Bartram, J. & Bentley, M. E. 2015 Drinking water quality governance: A comparative case study of Brazil, Ecuador, and Malawi. *Environmental Science and Policy* **48**, 186–195. https://doi.org/10.1016/j.envsci.2014.12.019.

Kim, B. R., Shin, J., Guevarra, R., Lee, J. H., Kim, D. W., Seol, K. H., Lee, J. H., Kim, H. B. & Isaacson, R. 2017 Deciphering diversity indices for a better understanding of microbial communities. *Journal of Microbiology and Biotechnology* **27**, 2089–2093. https://doi.org/10.4014/jmb.1709.09027.

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M. & Glöckner, F. O. 2013 Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Research* **41**, 1–11. https://doi.org/10.1093/nar/gks808.

Knee, K. & Encalada, A. 2013 Land use and water quality in rural cloud forest region (Intag, Ecuador). *River Research and Applications* **30**, 132–133. https://doi.org/10.1002/rra.

Kormas, K. A., Neofitou, C., Pachiadaki, M. & Koufostathi, E. 2010 Changes of the bacterial assemblages throughout an urban drinking water distribution system. *Environmental Monitoring and Assessment* **165**, 27–38. https://doi.org/10.1007/s10661-009-0924-7.

Koskey, A. M., Fisher, J. C., Eren, A. M., Ponce-Terashima, R., Reis, M. G., Blanton, R. E. & Mclellan, S. L. 2014 *Blautia* and *Prevotella* sequences distinguish human and animal fecal pollution in Brazil surface waters. *Environmental Microbiology Reports* **6**, 696–704. https://doi.org/10.1111/1758-2229.12189.

Krause, D. O., Dalrymple, B. P., Smith, W. J., Mackie, R. I. & McSweeney, C. S. 1999 16S rDNA sequencing of *Ruminococcus albus* and *Ruminococcus flavefaciens*: Design of a signature probe and its application in adult sheep. *Microbiology* **145**, 1797–1807. https://doi.org/10.1099/13500872-145-7-1797.

Lamb, P. D., Hunter, E., Pinnegar, J. K., Creer, S., Davies, R. G. & Taylor, M. I. 2019 How quantitative is metabarcoding: A meta-analytical approach. *Molecular Ecology* **28**, 420–430. https://doi.org/10.1111/mec.14920.

Legendre, P. & Legendre, L. 2012 *Numerical Ecology*. Elsevier, Amsterdam.

Legendre, P., Oksanen, J. & ter Braak, C. J. F. 2011 Testing the significance of canonical axes in redundancy analysis. *Methods in Ecology and Evolution* **2**, 269–277. https://doi.org/10.1111/j.2041-210X.2010.00078.x.

Levy, K., Nelson, K. L., Hubbard, A. & Eisenberg, J. N. S. 2008 Following the water: A controlled study of drinking water storage in northern coastal Ecuador. *Environmental Health Perspectives* **116**, 1533–1540. https://doi.org/10.1289/ehp.11296.

Levy, K., Hubbard, A. E., Nelson, K. L. & Eisenberg, J. N. S. 2009 Drivers of water quality variability in northern coastal Ecuador. *Environmental Science and Technology* **43**, 1788–1797. https://doi.org/10.1021/es8022545.

Levy, K., Nelson, K. L., Hubbard, A. & Eisenberg, J. N. S. 2012 Rethinking indicators of microbial drinking water quality for health studies in tropical developing countries: Case study in northern coastal Ecuador. *American Journal of Tropical Medicine and Hygiene* **86**, 499–507. https://doi.org/10.4269/ajtmh.2012.11-0263.

Li, C., Ling, F., Zhang, M., Liu, W. T., Li, Y. & Liu, W. 2017a Characterization of bacterial community dynamics in a full-scale drinking water treatment plant. *Journal of Environmental Science* **51**, 21–30. https://doi.10.1016/j.jes.2016.05.042.

Li, Q., Yu, S., Li, L., Liu, G., Gu, Z., Liu, M., Liu, Z., Ye, Y., Xia, Q. & Ren, L. 2017b Microbial communities shaped by treatment processes in a drinking water treatment plant and their contribution and threat to drinking water safety. *Frontiers in Microbiology* **8**, 2465. https://doi.org/10.3389/fmicb.2017.02465.

Li, F., Peng, Y., Fang, W., Altermatt, F., Xie, Y., Yang, J. & Zhang, X. 2018 Application of environmental DNA metabarcoding for predicting anthropogenic pollution in rivers. *Environmental Science & Technology* **52** (20), 11708–11719. https://doi.org/10.1021/acs.est.8b03869.

Maurice, L., López, F., Becerra, S., Jamhoury, H., Le Menach, K., Dévier, M. H., Budzinski, H., Prunier, J., Juteau-Martineau, G., Ochoa-Herrera, V., Quiroga, D. & Schreck, E. 2019 Drinking water quality in areas impacted by oil activities in Ecuador: Associated health risks and social perception of human exposure. *Science of the Total Environment* **690**, 1203–1217. https://doi.org/10.1016/j.scitotenv.2019.07.089.

Maynard, C. L., 2023 The microbiota in immunity and inflammation. In: *Clinical Immunology*, 6th edn. (Rich, R. R., Fleisher, T. A., Schroeder, H. W., Weyand, C. M., Corry, D. B. & Puck, J. M. eds.). Elsevier, Berkeley, CA, pp. 281–293. https://doi.org/10.1016/B978-0-7020-8165-1.00022.

McElroy, M., Dressler, T., Titcomb, G., Wilson, E., Deiner, K., Dudley, T., Eliason, E., Evans, N., Gaines, S., Lafferty, K., Lamberti, G., Li, Y., Lodge, D., Love, M., Mahon, A., Pfrender, M., Renshaw, M., Selkoe, K. & Jerde, C. 2020 Calibrating environmental DNA metabarcoding to conventional surveys for measuring fish species richness. *Frontiers in Ecology and Evolution* **8**, 276. https://doi.org/10.3389/fevo.2020.00276.

McLellan, S. L. & Eren, A. M. 2014 Discovering new indicators of fecal pollution. *Trends in Microbiology* **22** (12), 697–706. https://doi.org/10.1016/j.tim.2014.08.002.

Mclellan, S. L., Newton, R. J., Vandewalle, J. L., Shanks, O. C., Huse, S. M., Eren, A. M. & Sogin, M. L. 2013 Sewage reflects the distribution of human faecal Lachnospiraceae. *Environmental Microbiology* **15**, 2213–2227. https://doi.org/10.1111/1462-2920.12092.

McMurdie, P. J. & Holmes, S. 2013 Phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**. https://doi.org/10.1371/journal.pone.0061217.

Mideros, M. A. 2012 Ecuador: Defining and measuring multidimensional poverty, 2006–2010. *CEPAL Review* 49–67. https://doi.org/10.18356/9ed3c0a9-en.

Moreno, L., Pozo, M., Vancraeynest, K., Bain, R., Palacios, J. & Jácome, F. 2020 Integrating water-quality analysis in national household surveys: Water and sanitation sector learnings of Ecuador. *npj Clean Water* **3**, 23. https://doi.org/10.1038/s41545-020-0070-x.

Narciso-da-Rocha, C., Vaz-Moreira, I. & Manaia, C. M. 2014 Genotypic diversity and antibiotic resistance in Sphingomonadaceae isolated from hospital tap water. *Science of the Total Environment* **466–467**, 127–135. https://doi.org/10.1016/j.scitotenv.2013.06.109.

Newton, R. J., VandeWalle, J. L., Borchardt, M. A., Gorelick, M. H. & McLellan, S. L. 2011 Lachnospiraceae and bacteroidales alternative fecal indicators reveal chronic human sewage contamination in an urban harbor. *Applied and Environmental Microbiology* **77**, 6972–6981. https://doi.org/10.1128/AEM.05480-11.

Nie, K., Ma, K., Luo, W., Shen, Z., Yang, Z., Xiao, M., Tong, T., Yang, Y. & Wang, X. 2021 *Roseburia intestinalis*: A beneficial gut organism from the discoveries in genus and species. *Frontiers in Cellular and Infection Microbiology* **11**, 757718. https://doi.org/10.3389/fcimb.2021.757718.

Nierychlo, M., Andersen, K. S., Xu, Y., Green, N., Jiang, C., Albertsen, M., Dueholm, M. S. & Nielsen, P. H. 2020 MiDAS 3: An ecosystem-specific reference database, taxonomy and knowledge platform for activated sludge and anaerobic digesters reveals species-level microbiome composition of activated sludge. *Water Research* **182**, 115955. https://doi.org/10.1016/j.watres.2020.115955.

Norma Técnica Ecuatoriana NTE INEN 11:08. 2020 6th edn. Instituto Ecuatoriano de Normalización, Quito, Ecuador.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P., O'Hara, R. B., Simpson, G., Solymos, P., Stevens, M. H. H. & Wagner, H. 2012 *vegan: Community Ecology Package. R Package Version 2.0-3*. R Foundation for Statistical Computing, Vienna, Austria. CRAN. R-project. org/package = vegan.

Ortega-Paredes, D., Barba, P., Mena-López, S., Espinel, N., Crespo, V. & Zurita, J. 2019 High quantities of multidrug-resistant *Escherichia coli* are present in the Machángara urban river in Quito, Ecuador. *Journal of Water and Health* **18**, 67–76. https://doi.org/10.2166/wh.2019.195.

Paradis, E. & Schliep, K. 2019 Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528. https://doi.org/10.1093/bioinformatics/bty633.

Pinar-Méndez, A., Wangensteen, O. S., Præbel, K., Galofré, B., Méndez, J., Blanch, A. R. & García-Aljaro, C. 2022 Monitoring bacterial community dynamics in a drinking water treatment plant: An integrative approach using metabarcoding and microbial indicators in large water volumes. *Water* **14**, 1435. https://doi.org/10.3390/w14091435.

Ponce-Terashima, R., Koskey, A. M., Reis, M. G., McLellan, S. L. & Blanton, R. E. 2014 Sources and distribution of surface water fecal contamination and prevalence of schistosomiasis in a Brazilian village. *PLoS Neglected Tropical Diseases* **8**. https://doi.org/10.1371/journal.pntd.0003186.

Pozo, M., Serrano, J. C., Castillo, R. & Moreno, L. 2016 *Diagnóstico de los Indicadores ODS de Agua, Saneamiento E Higiene en Ecuador. ENEMDU 2016*. Estudios Temáticos – INEC, Quito, pp. 1–27

Rao, G., Eisenberg, J. N. S., Kleinbaum, D. G., Cevallos, W., Trueba, G. & Levy, K. 2015 Spatial variability of *Escherichia coli* in rivers of northern coastal Ecuador. *Water* **7**, 818–832. https://doi.org/10.3390/w7020818.

R Core Team. 2020 R: a language and environment for statistical computing.

Robinson, A., Wilde, J. & Allen-Vercoe, E., 2020 Fusobacteria: Physiology, form, and function. In: *Colorectal Neoplasia and the Colorectal Microbiome* (Floch, M. H., ed.). Academic Press, London. https://10.1016/B978-0-12-819672-4.00006-4.

Rompré, A., Servais, P., Baudart, J., De-Roubin, M. R. & Laurent, P. 2002 Detection and enumeration of coliforms in drinking water: Current methods and emerging approaches. *Journal of Microbiological Methods* **49**, 31–54. https://doi.org/10.1016/s0167-7012(01)00351-7.

Saxena, G., Bharagava, R. N., Kaithwas, G. & Raj, A. 2014 Microbial indicators, pathogens and methods for their monitoring in water environment. *Journal of Water and Health* **13**, 319–339. https://doi.org/10.2166/wh.2014.275.

Schliep, K. P. 2011 Phangorn: Phylogenetic analysis in R. *Bioinformatics* **27**, 592–593. https://doi.org/10.1093/bioinformatics/btq706.

Shelton, A. O., Gold, Z. J., Jensen, A. J., D Agnese, E., Andruszkiewicz, A. E., Van Cise, A., Gallego, R., Ramón-Laca, A., Garber-Yonts, M., Parsons, K. & Kelly, R. P. 2023 Toward quantitative metabarcoding. *Ecology* **104**, e3906. https://doi.org/10.1002/ecy.3906.

Shen, Z. 2016 Identification of host-specific genetic markers within 16S rDNA intervening sequences of 73 genera of fecal bacteria. *Journal of Data Mining in Genomics & Proteomics* **07**, 1–9. https://doi.org/10.4172/2153-0602.1000186.

Shen, Z., Duan, C., Zhang, C., Carson, A., Xu, D. & Zheng, G. 2013 Using an intervening sequence of *Faecalibacterium* 16S rDNA to identify poultry feces. *Water Research* **47**, 6415–6422. https://doi.org/10.1016/j.watres.2013.08.013.

Skelton, J., Cauvin, A. & Hunter, M. 2022 Environmental DNA metabarcoding read numbers and their variability predict species abundance, but weakly in non-dominant species. *Environmental DNA* **5**. https://doi.org/10.1002/edn3.355.

Stanish, L. F., Hull, N. M., Robertson, C. E., Harris, J. K., Stevens, M. J., Spear, J. R. & Pace, N. R. 2016 Factors influencing bacterial diversity and community composition in municipal drinking waters in the Ohio River basin, USA. *PLoS One* **11** (6), e0157966. https://doi.org/10.1371/journal.pone.0157966.

Sun, W., Liu, W., Cui, L., Zhang, M. & Wang, B. 2013 Characterization and identification of a chlorine-resistant bacterium, Sphingomonas TS001, from a model drinking water distribution system. *Science of the Total Environment* **458–460**, 169–175. https://doi.org/10.1016/j.scitotenv.2013.04.030.

Vavourakis, C. D., Heijnen, L., Peters, M. C. F. M., Marang, L., Ketelaars, H. A. M. & Hijnen, W. A. M. 2020 Spatial and temporal dynamics in attached and suspended bacterial communities in three drinking water distribution systems with variable biological stability. *Environmental Science & Technology* **54**, 14535–14546. https://doi.org/10.1021/acs.est.0c04532.

Vaz-Moreira, I., Nunes, O. C. & Manaia, C. M. 2011 Diversity and antibiotic resistance patterns of sphingomonadaceae isolates from drinking water. *Applied and Environmental Microbiology* **77**, 5697–5706. https://doi.org/10.1128/AEM.00579-11.

Villa-Achupallas, M., Rosado, D., Aguilar, S. & Galindo-Riaño, M. D. 2018 Water quality in the tropical Andes hotspot: The Yacuambi river (southeastern Ecuador). *Science of the Total Environment* **633**, 50–58. https://doi.org/10.1016/j.scitotenv.2018.03.165.

Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. 2007 Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology* **73**, 5261–5267. https://doi.org/10.1128/AEM.00062-07.

Wéry, N., Monteil, C., Pourcher, A. M. & Godon, J. J. 2010 Human-specific fecal bacteria in wastewater treatment plant effluents. *Water Research* **44**, 1873–1883. https://doi.org/10.1016/j.watres.2009.11.027.

WHO 2017 *Progress on Drinking Water, Sanitation and Hygiene: 2017 Update and SDG Baselines*, Report WHO/UNICEF, Geneva, Switzerland, Licence: CC BY-NC-SA 3.0 IGO, https://doi.org/10.1111/tmi.12329.

Zanacic, E., McMartin, D. W. & Stavrinides, J. 2017 From source to filter: Changes in bacterial community composition during potable water treatment. *Canadian Journal of Microbiology* **63**, 546–558. https://doi.org/10.1139/cjm-2017-0077.

Zemb, O., Achard, C. S., Hamelin, J., De Almeida, M. L., Gabinaud, B., Cauquil, L., Verschuren, L. M. G. & Godon, J. J. 2020 Absolute quantitation of microbes using 16S rRNA gene metabarcoding: A rapid normalization of relative abundances by quantitative PCR targeting a 16S rRNA gene spike-in standard. *MicrobiologyOpen* **9**, e977. https://doi.org/10.1002/mbo3.977.

Zeng, D. N., Fan, Z. Y., Chi, L., Wang, X., Qu, W. D. & Quan, Z. X. 2013 Analysis of the bacterial communities associated with different drinking water treatment processes. *World Journal of Microbiology and Biotechnology* **29**, 157315–84. https://doi.10.1007/s11274-013-1321-5.

Zhang, Y., Oh, S. & Liu, W. T. 2017 Impact of drinking water treatment and distribution on the microbiome continuum: An ecological disturbance's perspective. *Environmental Microbiology* **19**, 3163–3174. https://doi.org/10.1111/1462-2920.13800.

First received 22 September 2023; accepted in revised form 23 January 2024. Available online 8 February 2024